

Ratio: 5.249 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-525-867-1 x BC001715 ..
Align seg 1/1 to: BC001715 from: 1 to: 794

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
6 ATGGCGGTGCTGTACGTCCTGGCTGGCGGCTTCGGATCCTTGCTCT 55
|||||
17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
56 GGCTCCAGCGTGGCGGCTGTGACGACGAGGTGCTCATCAGAGCG 105
|||||
34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
106 TGCCACCGCATGGCCCAAGCAGCAGCCAGCTGCTGCTGCCAAGGCCAG 155
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
156 GCGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGTGGCCAAAGCT 205
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrp 84
|||||
206 GGATGACCTCGTCAACTGGCGCGCGGAGTCTCTGTGGCCCATGACCT 255
|||||
84 heGlyLeuAlaCysCysAlaValAlaGluMetMetHisMetAlaAlaProArg 100
|||||
256 TGGGCTGGCTGCTGCGCGCGGAGTATGATGATGATGATGATGATGATG 305
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
306 TACGACATGGACCGCTTTGGCGTGTCTTCCGCGCCAGCGCGCCAGTC 355
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
356 CACGCTGATGATCGTGGCGCGGCACACTCACCACCAAGATGCCCGCAGC 405
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
406 TTCGCAAGGCTCTACACACAGATCGCGGCGCGCTACGTTGCTCTCCATG 455
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValva 167
|||||
456 GGGAGCTGGCCACGAGGAGGCTACTACCACTATTCTACTCGGTGT 505
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
506 GAGGGGCTGCGACCGATCGTGGCGGTGACATCTACATCCAGGCTGCC 555
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
556 CACCTACGCGCGAGCGCTGCTACGCGCATCCTGCGAGCTCGACAGGAAG 605
|||||
201 IleLysArgGluArgAlaLeuGlnIleTyrTyrArgArg 213
|||||
606 ATCAAGCGGACGCGAGGCTCGAGATCTGTACCGCAGG 644

seq_name: gb_pr:BC005954
seq_documentation_block:
LOCUS BC005954 787 bp mRNA PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:14592 IMAGE:4276489, mRNA, complete cds.
ACCESSION BC005954
VERSION BC005954.1 GI:13543602
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 787)

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ARCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxli@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 21 Row: k Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
Location/Qualifiers
1..787
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:14592 IMAGE:4276489"
/tissue_type="brain, primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
21..662
/codon_start=1
/product="Unknown (protein for MGC:14592)"
/protein_id="AAH05954.1"
/db_xref="GI:13543603"
/translation="MAVLSAPGLRFRILGLSSVGLAVQARGVHQSVATDGPSTQ
PDKARAVAPKPSRGEYVAKLDLVMNRKSLWPMFTGLACCAVEMHMAAPRYD
MDRFVVFRRASPRQSDVIVAGLTNKNAPALRKYVQDMPFRYVYVSMGSCANGGYY
HYSYSVVRGCDRIVPVDIYIPGCPPTAEALLYILQLOKIKRRLQIWR"

BASE COUNT 152 a 262 c 244 g 129 t
ORIGIN

alignment_scores:
Quality: 1108.00 Length: 213
Ratio: 5.226 Gaps: 0
Percent Similarity: 99.531 Percent Identity: 99.531

alignment_block:
US-09-525-867-1 x BC005954 ..
Align seg 1/1 to: BC005954 from: 1 to: 787

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
21 ATGGCGGTGCTGTACGTCCTGGCTGGCGGCTTCGGATCCTTGCTCT 70
|||||
17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
71 GGCTCCAGCGTGGCGGCTGTGACGACGAGGTGCTCATCAGAGCG 120
|||||
34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
121 TGCCACCGCATGGCCCAAGCAGCAGCCAGCTGCTGCTGCCAAGGCCAG 170
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
171 GCGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGTGGCCAAAGCT 220
|||||

```

67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
221 GATGACCTCGTCAACTGGGCGCGGAGTTCTCTGTGGCCATGACCT 270
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
271 TCGGCTGGCCCTGCTGGCGCTGGAGATGATGCATGGCAGCAGCCCGC 320
101 TyrAspMetAspArgPheGlyValValPheAlaSerProArgGlnSe 117
321 TAGGACATGACCGCTTTGGCGTGGTCTTCCGCGCAGCGCGCCAGTC 370
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
371 CGAGCTCATGATCGTGGCGCGGACACTACCAACAAGATGGCCCGCAGC 420
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
421 TTCGCAAGGTCTACGACCATGCGGAGCGCGCTACGTGGTCTCCATG 470
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
471 GAGAGCTGGCCCAACGAGGAGGCTACTACCACTATTCTCTACTCGTGGT 520
167 IaArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
521 GAGGGCTGGCAGCCGATCGTGGCGGTGACATCTACATCCAGGCTGCC 570
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuArgLys 200
571 CACCTACGGCGAGGCGCTGCTCTACGGCATCTGCAGTGCAGAGGAAG 620
201 IleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
621 ATCAAGCGGAGCGGAGGCTGACATCTGGTACCGCAGG 659

```

seq_name: gb_pr:AF060512

```

seq_documentation_block:
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Mao,Y.M., Xie,Y., Huang,X.Y., Ying,X. and Dai,J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China

```

FEATURES

Location/Qualifiers

```

1..754
/organism="Homo sapiens"
/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..654
/codon_start=1

```

```

/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..654
/codon_start=1

```

```

CDS
120 a 267 c 239 g 128 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
Quality: 1046.50 Length: 215
Ratio: 4.960 Gaps: 2
Percent Similarity: 98.140 Percent Identity: 97.209
alignment_block:
US-09-525-867-1 x AF060512
Align seg 1/1 to: AF060512 from: 1 to: 754
1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
10 ATGGCGGTGCTGTACGCTCTGGCGCTGGCGGCTTCCGATCCTTGGTCT 59
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnServ 34
60 GCCTCTCAGCGTGGCGCTGGCTGTGAGGCTGTCAGGCTGTCCATCAGAGC 109
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
110 TGCCACCATGATGCCCCAAGCAGCAGCCAGCTGCCCTGCCAAGGCCAGA 159
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
160 GCGGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGTGGCCAAAGCT 209
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
210 GATGACCTCGTCAACTGGGCGCGGAGTTCTCTGTGGCCATGACCT 259
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
260 TCGGCTGTGGCTGTGGCGCTGGAGATGATGCACATGGCAGCAGCCCGC 309
101 TyrAspMetAspArgPheGlyValValPheAlaSerProArgGlnSe 117
310 TAGGACATGACCGCTTTGGCGTGGTCTTCCGCGCAGCGCGCCAGTCTC 359
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
360 CGAGCTCATGATCGTGGCGCGGACACTACCAACAAGATGGCCCGCAGC 409
134 euArgLys.Val...TyrAspGlnMetProGluProArgTyrValValSe 149
410 TTGCAAGTCTCCCCCAGCAGCAGATGCCGAGCGCGCTACGTGGTCTC 459
149 rMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrServ 166
460 CATGGGAGCTGGCCACAGCGGAGGCTACTACCACTATTCTCTACTGG 509
166 alValArgGlyCysAspArgIleValProValAspIleTyrIleProGly 182
510 TGTGAGGGCTCGAGCCGATCGTGGCGTGGACATCTACATCCAGGC 559
183 CysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnAr 199
560 TGCCCACTACGCGCAGGCGCTGTCTACGCGCATCTCTCAGCTGCAGTC 608
199 gLysIleLysArgGluArgGlnLeuIleTrpTyrArgArg 213
609 GAAGATCAAGCGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 651

```

seq_name: gb_on:BTPSSTS

seq_documentation_block:

```

LOCUS BTPSSTS 752 bp mRNA MAM
DEFINITION B.taurus mRNA for pSST subunit of the NADH: ubiquinone
oxidoreductase complex.
ACCESSION X65020
VERSION X65020.1 GI:11256
KEYWORDS NADH-ubiquinone oxidoreductase subunit.
SOURCE cow.

```

29-NOV-1994


```

BASE COUNT      8656 a   6105 c   6205 g   8597 t
ORIGIN

alignment_scores:
  Quality: 756.50      Length: 197
  Ratio: 4.274         Gaps: 3
  Percent Similarity: 89.848  Percent Identity: 71.574

alignment_block:
  US-09-525-867-1 x AC014169 ..

Align seg 1/1 to: AC014169 from: 1 to: 29563

18 ArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerVa 34
19629 CGCTCTGAATTGGCCCGCTGGT.....CGGCCACAAACAGGTTCCAG 19672

34 lAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 51
19673 TTGCCGAAGTGGCCCAAAATCTGCCCAAGAAGGCTA.....CTCT 19713

51 lavalAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
19714 CGGTTGGCCACCAACAGTCTCCGTTGCGGAGTGGTCACTGGCCAGACT 19763

67 uAspAspLeuValAsnTtpAlaArgArgSerSerLeuTyrProMetThrp 84
19764 GACGATGTCCTCACTGGGTCGCAAGGTCGATCTGGCCACTGACTT 19813

84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
19814 TCGGTTGGCCCTGCTGTCGCGTCGCAATCATGCACATCGCTGCTCGCGT 19863

101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
19864 TAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 19913

117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
19914 CGATGTCATCATCGTCGCTGGCCAGCTGACCAACAAATATGACCCGCGC 19963

134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
19964 TCGAAAGTCTTACGACCAATATGCCCCGACCCACGTTGGGTCACTCCATG 20013

151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
20014 GGCAGCTGTGCCAAGCGGGCGGCTACTACCATTAATCTACTCGTCTGT 20063

167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
20064 CGGTGGCTGCGATAGATAAATCCCGTCGACATATACGTACCCGCTGTC 20113

184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
20114 CGCCACCCCGGAGGCGCTCATCTACGCGGTTTTCAGCTGCAGAGAAG 20163

201 lIeLysArgGluArgLeuGlnIleTyrTyrArgArg 213
20164 GTTAAGCGATGAAGACGCTCCAGATGTTGGTATAGGAAG 20202

seq_name: gb_htg:AC010919

seq_documentation_block:
LOCUS      AC010919      140685 bp      DNA      13-DEC-1999
DEFINITION Drosophila melanogaster chromosome x clone BACR32K23 (D891) RPOI-98
            32.K.23 map 14A-14A strain y; cn bw sp. *** SEQUENCING IN PROGRESS
            *** 77 unordered pieces.
ACCESSION  AC010919
VERSION     AC010919.9  GI:563419
KEYWORDS   HTG; HTGS-PHASE1.
SOURCE     fruit fly.

```

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hickie,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Swirskas,R.R., Wan,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 140685)
Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,B., Sequeira,A., Sethi,H., Snir,E.,
Swirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (27-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Dec 13, 1999 this sequence version replaced gi:6449494.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 77 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 674: contig of 674 bp in length
* 675
754: gap of unknown length
* 755
1507: contig of 753 bp in length
* 1508
1587: gap of unknown length
* 1588
2889: contig of 1302 bp in length
* 2890
2969: gap of unknown length
* 2970
3554: contig of 585 bp in length
* 3555
3634: gap of unknown length
* 3635
4239: contig of 605 bp in length
* 4240
4319: gap of unknown length
* 4320
4698: contig of 379 bp in length
* 4699
4778: gap of unknown length
* 4779
5295: contig of 517 bp in length
* 5296
5375: gap of unknown length
* 5376
5935: contig of 560 bp in length
* 5936
6015: gap of unknown length
* 6016
6640: gap of unknown length
* 6641
7460: contig of 820 bp in length
* 7461
7540: gap of unknown length
* 7541
9170: contig of 1630 bp in length
* 9171
9250: gap of unknown length
* 9251
9882: contig of 632 bp in length
* 9883
9962: gap of unknown length
* 9963
11247: contig of 1284 bp in length
* 11248
11326: gap of unknown length
* 11327
12802: contig of 1476 bp in length
* 12803
12882: gap of unknown length
* 12883
14230: contig of 1348 bp in length
* 14231
14310: gap of unknown length
* 14311
15711: contig of 1401 bp in length


```
* 15712 15791: gap of unknown length
* 15792 16237: contig of 446 bp in length
* 16398 16317: gap of unknown length
* 16318 17991: contig of 1674 bp in length
* 17992 18071: gap of unknown length
* 18072 19191: contig of 1120 bp in length
* 19192 19271: gap of unknown length
* 19272 21324: contig of 2053 bp in length
* 21325 21404: gap of unknown length
* 21405 23898: contig of 2494 bp in length
* 23899 23979: gap of unknown length
* 23980 27079: contig of 3101 bp in length
* 27080 27160: gap of unknown length
* 27161 29602: contig of 2443 bp in length
* 29603 29682: gap of unknown length
* 29683 31919: contig of 2237 bp in length
* 31920 31999: gap of unknown length
* 32000 34542: contig of 2543 bp in length
* 34543 34622: gap of unknown length
* 34623 36445: contig of 1823 bp in length
* 36446 36525: gap of unknown length
* 36526 40018: contig of 3493 bp in length
* 40019 40098: gap of unknown length
* 40099 42061: contig of 1963 bp in length
* 42062 42141: gap of unknown length
* 42142 48164: contig of 6023 bp in length
* 48165 48244: gap of unknown length
* 48245 54296: contig of 6052 bp in length
* 54297 54376: gap of unknown length
* 54377 66627: contig of 12251 bp in length
* 66628 66707: gap of unknown length
* 66708 77734: contig of 11027 bp in length
* 77735 77814: gap of unknown length
* 77815 88534: contig of 10720 bp in length
* 88535 88614: gap of unknown length
* 88615 101020: contig of 12406 bp in length
* 101021 101100: gap of unknown length
* 101101 115583: contig of 14483 bp in length
* 115584 116236: contig of 573 bp in length
* 116237 116317: gap of unknown length
* 116318 116809: contig of 493 bp in length
* 116810 116889: gap of unknown length
* 116890 117344: contig of 454 bp in length
* 117345 117423: gap of unknown length
* 117424 118337: contig of 914 bp in length
* 118338 118417: gap of unknown length
* 118418 118910: contig of 493 bp in length
* 118911 118990: gap of unknown length
* 118991 119603: contig of 613 bp in length
* 119604 119684: gap of unknown length
* 119685 120333: contig of 650 bp in length
* 120334 120413: gap of unknown length
* 120414 120968: contig of 555 bp in length
* 120969 121048: gap of unknown length
* 121049 121685: contig of 637 bp in length
* 121686 121755: gap of unknown length
* 121756 122379: contig of 614 bp in length
* 122380 122459: gap of unknown length
* 122460 122949: contig of 490 bp in length
* 122950 123029: gap of unknown length
* 123030 123632: contig of 603 bp in length
* 123633 123712: gap of unknown length
* 123713 124146: contig of 434 bp in length
* 124147 124266: gap of unknown length
* 124267 124732: contig of 506 bp in length
* 124733 124812: gap of unknown length
* 124813 125420: contig of 608 bp in length
* 125421 125500: gap of unknown length
* 125501 126104: contig of 604 bp in length
* 126105 126184: gap of unknown length
* 126185 126663: contig of 479 bp in length
* 126664 126743: gap of unknown length

* 126744 127183: contig of 440 bp in length
* 127184 127863: gap of unknown length
* 127864 127903: contig of 540 bp in length
* 127904 127883: gap of unknown length
* 127884 128725: contig of 842 bp in length
* 128726 128805: gap of unknown length
* 128806 129167: contig of 362 bp in length
* 129168 129247: gap of unknown length
* 129248 129335: contig of 688 bp in length
* 129336 130015: gap of unknown length
* 130016 130560: contig of 545 bp in length
* 130561 130640: gap of unknown length
* 130641 131152: contig of 512 bp in length
* 131153 131232: gap of unknown length
* 131233 131826: contig of 594 bp in length
* 131827 132383: contig of 477 bp in length
* 132384 132463: gap of unknown length
* 132464 133016: contig of 553 bp in length
* 133017 133096: gap of unknown length
* 133097 133747: contig of 651 bp in length
* 133748 133827: gap of unknown length
* 133828 134276: contig of 449 bp in length
* 134277 134356: gap of unknown length
* 134357 134964: contig of 608 bp in length
* 134965 135044: gap of unknown length
* 135045 135383: contig of 339 bp in length
* 135384 135463: gap of unknown length
* 135464 136015: contig of 552 bp in length
* 136016 136095: gap of unknown length
* 136096 136547: contig of 452 bp in length
* 136548 136627: gap of unknown length
* 136628 137067: contig of 440 bp in length
* 137068 137147: gap of unknown length
* 137148 137600: contig of 453 bp in length
* 137601 137680: gap of unknown length
* 137681 138054: contig of 374 bp in length
* 138055 138134: gap of unknown length
* 138135 138352: contig of 218 bp in length
* 138353 138432: gap of unknown length
* 138433 138859: contig of 427 bp in length
* 138860 138939: gap of unknown length
* 138940 139149: contig of 210 bp in length
* 139150 139229: gap of unknown length
* 139230 139727: contig of 498 bp in length
* 139728 139807: gap of unknown length

alignment_scores:
Quality: 756.50 Length: 197
Ratio: 4.274 Gaps: 3
Percent Similarity: 89.848 Percent Identity: 71.574

alignment_block:
US-09-525-867-1 x AC010919 ..
Align seg 1/1 to: AC010919 from: 1 to: 140685

18 ArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerVa 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103611 CGCTGTGAATTGGCCCTGGT.....GCGCCAAACAACACCTTCAG 103654

34 lAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgA 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103655 TTGCCGAAGTGGCCCAAAATCTGCCAAGAGGGCTA.....CTCT 103695

51 laVal.lAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103696 CCGTTTCGCACCAACACAGTCTCGGTTTCGGAGTGGTCACTGCCACACT 103745

67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103746 GGACGATCTGCTCAACTGGGTGCGCAAGGGCTCGATCTGCCACTGACTT 103795
```


27797	27876	gap of unknown length
27877	28981	contig of 1105 bp in length
27882	29061	gap of unknown length
27882	29070	contig of 909 bp in length
29062	30050	gap of unknown length
29071	31555	contig of 1505 bp in length
30051	31635	gap of unknown length
31556	33133	contig of 1488 bp in length
31636	33203	gap of unknown length
33124	34199	contig of 996 bp in length
33204	34279	gap of unknown length
34200	35179	contig of 900 bp in length
34280	35259	gap of unknown length
35180	36555	contig of 1296 bp in length
35260	36635	gap of unknown length
36556	37489	contig of 854 bp in length
36636	37569	gap of unknown length
37490	39224	contig of 1655 bp in length
39225	39304	gap of unknown length
39305	41596	contig of 2292 bp in length
41597	41676	gap of unknown length
41677	43339	contig of 1663 bp in length
43340	43419	gap of unknown length
43420	45087	contig of 1668 bp in length
45088	45167	gap of unknown length
45168	46523	contig of 1356 bp in length
46524	46603	gap of unknown length
46604	48539	contig of 1956 bp in length
48560	48639	gap of unknown length
48640	50589	contig of 1950 bp in length
50590	50669	gap of unknown length
50670	52607	contig of 1938 bp in length
52608	52687	gap of unknown length
52688	53881	contig of 1194 bp in length
53882	53961	gap of unknown length
53962	55649	contig of 1688 bp in length
55650	55729	gap of unknown length
55730	57438	contig of 1709 bp in length
57439	57518	gap of unknown length
57519	59734	contig of 2216 bp in length
59735	59814	gap of unknown length
59815	62611	contig of 2807 bp in length
62622	62701	gap of unknown length
62702	65239	contig of 2538 bp in length
65240	65319	gap of unknown length
65320	68682	contig of 3363 bp in length
68683	68762	gap of unknown length
68763	70859	contig of 2097 bp in length
70860	70939	gap of unknown length
70940	73115	contig of 2176 bp in length
73116	73195	gap of unknown length
73196	75979	contig of 2784 bp in length
75980	76059	gap of unknown length
76060	78751	contig of 2691 bp in length
78752	78830	gap of unknown length
78831	81983	contig of 3153 bp in length
81984	82063	gap of unknown length
82064	84253	contig of 2190 bp in length
84254	84333	gap of unknown length
84334	88351	contig of 4018 bp in length
88352	88431	gap of unknown length
88432	92028	contig of 3597 bp in length
92029	92108	gap of unknown length
92109	96970	contig of 4862 bp in length
96971	97050	gap of unknown length
97051	102071	contig of 5021 bp in length
102072	102151	gap of unknown length
102152	108106	contig of 5955 bp in length
108107	108186	gap of unknown length
108187	113640	contig of 5454 bp in length
113641	113720	gap of unknown length
113721	120433	contig of 6723 bp in length
120434	120543	gap of unknown length

59157 GTTAAAGGATGAAGACGCTCCAGATGTTGGTATAGGAAG 59195

seq_name: gb_in:AE003500

seq_documentation_block:

LOCUS AE003500 327446 bp DNA INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386053 section 17
of 30, complete sequence.

ACCESSION AE003500 AE002593

VERSION AE003500.2 GI:10728273

KEYWORDS HTG.

SOURCE fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 327446)

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,

Amantides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,

George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,

Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,O., Chen,L.X.,

Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,

Wap,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor

Milos,G.L., Abril,J.F., Agbayani,A., An,H.J.,

Andrews-Frankkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,

Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,

Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,

Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,

Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,

Cherry,J.M., Cawley,S., Dahike,C., Davenport,L.B., Davies,P., de

Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,

Dodson,K., Doup,I.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,

Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieria,S.,

Flisbachmann,W., Foster,K., Gabriellian,A.E., Garg,N.S.,

Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,

Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J.,

Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,

Wei,M.H., Ibegwam,C., Jaitli,M., Kalush,F., Karpen,G.H., Ke,Z.,

Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,

Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,

Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattel,B., McIntosh,T.C.,

McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,

Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,

Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,

Nusskern,D.R., Paclet,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,

Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,

Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,

Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,

Stapleton,M., Strong,R., Sun,E., Svirska,R., Tector,C., Turner,R.,

Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A.,

Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T.,

Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yen,R.F.,

Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,

Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,

Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

The genome sequence of *Drosophila melanogaster*

Science 287 (5461), 2185-2195 (2000)

20196006

2 (bases 1 to 327446)

Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.

Direct Submission

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,

Rockville, MD, USA

On Oct 9, 2000 this sequence version replaced gi:7293108.

Location/Qualifiers

1. 327446

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/chromosome="X"

complement(join(772. 903,968. 1504,1572. 2077,2153. 2821,

4145. >4380))

/gene="CG9281"

/product="CT26402"

mrna

gene

/db_xref="FLYBASE:FBan0009281"

/db_xref="FLYBASE:FBgn0030672"

complement(<772. >4380)

/gene="CG9281"

/map="13E12-13E14"

/db_xref="FLYBASE:FBan0009281"

/db_xref="FLYBASE:FBgn0030672"

complement(join(772. 903,968. 1504,1572. 2077,2153. 2813))

/gene="CG9281"

/note="CG9281 gene product"

/codon_start=1

/db_xref="FLYBASE:FBan0009281"

/db_xref="FLYBASE:FBgn0030672"

/protein_id="AAP48493.1"

/db_xref="GI:7293109"

/translation="MSRAKKREAAKKKDKARKKIQIIPSTKKANCVPERELTEBKL

CAKLEEARISAEARCTSLAVHPSRDVYKIANFSITFFGSELLQDMLNLCGRRY

GLGLNGCGSSLEDAQEITDIYERLDDMSADIAEVAARILHGLGFRKMOQKQAD

EKLAEELMSGEEDAEQITDIYERLDDMSADIAEVAARILHGLGFRKMOQKQAD

FSGWRMRIALARLVKPHLLLDPTNHLDDACVLEELKTYKRILVLIHQSD

FLNGVCTNIHLTKGLKTYGNYEAFVTRMELLENQMKQYNWQDOISHMKNIAR

FGHSDAKLAAQOSKBTAKMAVQGLTEKVSDDKVLNFYFPCGKQVPPVIMQVNH

FRYDEPPIYKLESGIDLDTPALVGPNGAGKSTLLKLLYGLDLPVPTSGMIRKNSH

RIARYHOLHELLEDLQASPLEYMMRAFPPYKEEEMRKIIIGRYGLIGQQVCPIQLS

DGQRCRVFWALWQVPHLLLDPTNHLDDACVLEELKTYKRILVLIHQSD

NOVAEEIWOCEKETYTKWGGIDLYDKHLKNKITSNERKKAKAG"

join(<4963. 5295,5348. >5848)

/gene="CG15601"

/product="CT35724"

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

<4963. >5848

/gene="CG15601"

/map="13E14-13E14"

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

join(4963. 5295,5348. 5848)

/gene="CG15601"

/note="CG15601 gene product"

/codon_start=1

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

/protein_id="AAF48494.1"

/db_xref="GI:7293110"

/translation="MSKLSDTKIPTEFLEAYRRQPCLYNTLDSYKNRVSREAYGA

IIRSLKIPQTVSDIKLIKSVTVYKELRWREKELGRTYEPKLPFLRLADGFLR

SVSLSHCKRQGNKNSDAOLTTIKSDTKLCTAAADITMSDALDEEAEVNGEFLR

ECPLESRPTASICKDSTLCLADQPCQHYSGQSSQQLPHTMAQRKSKYITSLDS

AGEDDLIFQCSIAQURTIPDSISYKVLRIQVLEAFEGQGFQSTEVNTQLQNT

F"

join(<7954. 8220,8556. 8795,8865. 9086,9145. 9416,

9497. 9993,10131. 10515,10567. 15238,15758. 18218,

18282. 18755,18814. 19154,19217. 20287,20378. 20513,

20573. 22032,22060. 23197,23270. 23671,23739. 23927,

23992. 24313,24377. 24542,24617. 24807,24885. >24987)

/gene="CG8184"

/product="CT24358"

/db_xref="FLYBASE:FBan0008184"

/db_xref="FLYBASE:FBgn0030674"

<7954. >24987

/gene="CG8184"

/map="13E14-13E16"

/db_xref="FLYBASE:FBan0008184"

/db_xref="FLYBASE:FBgn0030674"

join(7954. 8220,8556. 8795,8865. 9086,9145. 9416,

9497. 9993,10131. 10515,10567. 15238,15758. 18218,

18282. 18755,18814. 19154,19217. 20287,20378. 20513,

20573. 22032,22060. 23197,23270. 23671,23739. 23927,

23992. 24313,24377. 24542,24617. 24807,24885. 24987)

/gene="CG8184"

source

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES


```

BASE COUNT      23093 a 18268 c 19032 g 23214 t
ORIGIN
alignment_scores:
  Quality: 743.00      Length: 154
  Ratio: 4.921         Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766
alignment_block:
US-09-525-867-1 x AC019889/rev ..
Align seg 1/1 to reverse of: AC019889 from: 1 to: 83607
60 GlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaArgAr 76
14234 GCGAGTGGACCTGGCCGAGTGGATGACCTGCTAACTGGGACGCAA 14185
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGlu 93
14184 GGGATCCCTGGCCGCTGACCTTGGTCTGCTGCTGCTGCTGCTGCTG 14135
93 eMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
14134 TGATCACATTCGCGCTCTCGTACGACATGATGCTACGGTGTGGTG 14085
110 PheArgAlaSerProArgGlnSerAspValMetLeuValAlaGlyThrLe 126
14084 TTCCGAGCATTCCTCCGACGCGGACGTGCTCATCGTGGCCGGAACCT 14035
126 uThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
14034 GACCACACAGATGGCCGCGCTTCGGAACATCTACGACCATGCGCG 13985
143 LuProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
13984 AGCCGAGATGGGTCTATTCATTCGAGTGGSCAGTTCGCAATGTGTGG 13935
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgLeuValProVa 176
13934 TACCCTACTCTCTACTCGTGGTTCGCGGCTGGCATGTCATGTTCGGT 13885
176 LaspleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG 193
13884 GGACATCTACGTCGCGGATGTCGCCACCGCCGAGGCTTAATGTACG 13835
193 LylleuGlnLeuGlnArgLysLysLysArgGluArgArgLeuGlnIle 209
13834 GAATCTGAGCTGCAGAGAGAGGTAAAGCGCATGAGGACCCCTGCAGATG 13785
210 TrpTyrArgArg 213
13784 TGTACAGGAAG 13773
seq_name: gb_in:AC007856
seq_documentation_block:
LOCUS      AC007856.1 175124 bp DNA INV 17-FEB-2001
DEFINITION Drosophila melanogaster, chromosome 3R, region 99B-99B, BAC clone
ACCESSION  BACR05L03, complete sequence.
VERSION    AC007856.7 GI:12957650
KEYWORDS   HTG.
SOURCE     fruit fly.
ORGANISM  Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 175124)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,

```

```

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Scheeler,F.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 3R, region 99B-99B
Unpublished
2 (bases 1 to 175124)
Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhorf,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (17-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 17, 2001 this sequence version replaced gi:6633905.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (http://www.fruitfly.org/sequence/) or send email
to bgp@fruitfly.berkeley.edu.
FEATURES             source
     Location/Qualifiers
         1..175124
             /organism="Drosophila melanogaster"
             /strain="y; cn bw sp"
             /db_xref="taxon:7227"
             /chromosome="3R"
             /map="99B-99B"
             /clone="BACR05L03 (D795)"
             /clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial ECORI in
pBACE3.6)"
BASE COUNT      47809 a 39206 c 39325 g 48784 t
ORIGIN
alignment_scores:
  Quality: 743.00      Length: 154
  Ratio: 4.921         Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766
alignment_block:
US-09-525-867-1 x AC007856 ...
Align seg 1/1 to: AC007856 from: 1 to: 175124
60 GlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaArgAr 76
22222 GCGAGTGGACCTGGCCGAGTGGATGACCTGCTAACTGGGACGCAA 22271
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGlu 93
22272 GGGATCCCTGGCCGCTGACCTTGGTCTGCTGCTGCTGCTGCTGCTG 22321
93 eMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
22322 TGATCACATTCGCGCTCTCGTACGACATGATGATGATGATGATGATG 22371

```



```

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Ling,H.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Richards,S., Sequeira,A., Sethi,H.,
Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M.,
Zieran,L.L. and Rubin,G.M.
TITLE Sequencing of Drosophila chromosome 3R, region 99A7-99B5
JOURNAL Unpublished
2 (bases 1 to 311043)
REFERENCE
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Ling,H.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Richards,S., Sequeira,A., Sethi,H.,
Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M.,
Zieran,L.L. and Rubin,G.M.
TITLE Submitted
JOURNAL Direct Submission
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
COMMENT On Sep 4, 1999 this sequence version replaced gi:5713324
gi:5670685.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgp@fruitfly.berkeley.edu.
FEATURES
source
1..311043
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="99A7-99B5"
/clone="BACS BACR48K01 (D488) and BACR48O23 (D498)"
/clone_lib="RPCi-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"
/notes="BACR48K01 (D488) and BACR48O23 (D498) were
sequenced as a project. BACR48K01 extends from bp 1 to bp
183,934. BACR48O23 extends from bp 142,325 to bp
311,043."
BASE COUNT 87124 a 58661 c 68467 g 86791 t
ORIGIN

alignment_scores:
Quality: 743.00 Length: 154
Ratio: 4.921 Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766

alignment_block:
US-09-525-867-1 x AC009888 ..
Align seg 1/1 to: AC009888 from: 1 to: 311043

60 GlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaArgAr 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231343 GCGAGTGACCTGTGCCAGTGTGGATGACCTGCTAAACGGGACGCAA 231392

76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluM 93
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231393 GGGATCCCTCGCGCGTGCACCTTTGTCTCGCTGCTGGCGGTGAAA 231442

93 etMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231443 TGATGCACATTGGCGCTCTCGCTACGACATGGATCGCTACGGTGTGGTG 231492

110 PheArgAlaSerProArgGlnSerAspValMetIleValalaglyThrLe 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231493 TTCCGAGGATCTCTCGCGAGCGGAGCTGCTCATCTGCGCGAACCCT 231542

126 tThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231543 GACCAACAAGATGGACCGCGCTTTCGGAAGATCTACGACCATGGCCG 231592

143 luProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231593 AGCCGAGATGGGTCTATTTCCATGGGCAGTTTCGCCAATGGTGGCGGTAC 231642

160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231643 TACCACTACTCTACTCGTGTGTTTCGGCGCTGGCATGCGATTGTTCGGT 231692

176 lasPileTyrIleProGlyCysProThrAlaGluAlaLeuLeuTyrG 193
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231693 GGACATCTACGTCCCGGATGTCGCCACCGCGAGGCGCTTAATGTACG 231742

193 ylleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIle 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231743 GAATCCTGCAGTCGAGAAGAAGGTAAAGCGCATGAGGACCGCTGCAGATG 231792

210 TrpTyrArgArg 213
|||||:|||||:
231793 TGGTACAGGAAG 231804

seq_name: gb_pl:BONADHPST
seq_documentation_block:
LOCUS BONADHPST 911 bp mRNA PLN 28-SEP-1995
DEFINITION B.oleracea mRNA for PSST subunit of NADH:ubiquinone oxidoreductase.
ACCESSION X82274
VERSION X82274.1 GI:562281
KEYWORDS NADH-ubiquinone oxidoreductase; NADH:ubiquinone oxidoreductase
subunit PSST.
SOURCE
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 911)
Pogson,B.J., Downs,C.G., Davies,K.M., Morris,S.C. and
Buchanan-Wollaston,V.
Nucleotide sequence of a cDNA clone from broccoli with high
identity with the PSST subunit of NADH:ubiquinone oxidoreductase
Plant Physiol. 108 (2), 859-860 (1995)
9534517
2 (bases 1 to 911)
Pogson,B.J.
Direct Submission
TITLE Submitted (07-SEP-1994) B.J. Pogson, University of Arizona, Dept of
Plant Sciences, 303 Forbes Hall, Tucson 85721, Arizona, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..911
/organism="Brassica oleracea"
/cultivar="Shogun"
/db_xref="taxon:3712"
/dev_stage="mature plant"
/tissue_type="florets"
/clone="pBROC4"
/notes="L."
35..682
/codon_start=1
CDS

```

```

/product="PSST subunit of NADH: ubiquinone oxidoreductase"
/protein_id="CAA57725.1"
/db_xref="GI:562282"
/translation="MAMITNTATRLPLVQLSHRAAAVSHLHTSLPALSPATTPTSYT
YDLPFGILIRPSPQSDCMIVAGTLFNKAPALRKVDMPPEPRWVISMGCANGGG
YYHYSYVVRGCDRIVPVDIYVPGCPPTAEALLYGLLQKKINRRKDFLEWNNK"
BASE COUNT      226 a      234 c      218 g      233 t
ORIGIN
1000
alignment_scores:
  Quality: 736.00      Length: 181
  Ratio: 4.543        Gaps: 0
  Percent Similarity: 89.503      Percent Identity: 73.481
alignment_block:
US-09-525-867-1 x BONADHPST
Align seg 1/1 to: BONADHPST from: 1 to: 911
33 SerValAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAl 49
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
137 TCCCGCCGACACACCGAGCTTCATACACAGCAGGCTCCCTCCATCGAC 186
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 aArgAlaValAlaProLysProSerSerArgGlyGluTyrValValAlaL 66
187 CTCGGCTCCTCCCGGGTCTCTCGAGAGCGGGAGTTCGTGATCTCAA 236
66 yLeuAspLeuValAlaAsnTyrPalaArgSerSerLeuTyrProMet 82
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
237 AGGTAGATGATCTGATGAAGTACGCGCCGAGGAGGATGATCGGCCCATG 286
83 ThrPheGlyLeuAlaCysCysAlaValAlaGluMetMetHisMetAlaAlaPr 99
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
287 ACCITCGGGTCCCTGCTCGCGCGTGGAGATGATCATACAGGTCGCGC 336
99 oArgTyrAspMetAspArgPheGlyValValPheArgAlaSerProArgG 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 TCGCTACGATCTGGACCGGATTCGGTATCATCTTCAGGCGGAGTCTCGCC 386
116 lnSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaPro 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 AATCGGATGTATGATGTGCGCGGACGCTTACTAACAGATGCTGCC 436
133 AlaLeuArgLysValTyrAspGlnMetProGluProArgTyrValValSe 149
437 GCTCTTCGAAGTTTATGACAGATCCCGGAGCCCAAGGTGGTGATTC 486
149 rMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrServ 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 AATGGGAAGTTGTCCAACGAGGTGGATATTATCATCTACTCTACTCGG 536
166 alValArgGlyCysAspArgIleValProValAspIleTyrIleProGly 182
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
537 TCGTTCGAGGTGGACAGAAATGTGCGACGTGACATCATGCTCCCTGT 586
183 CysProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnAr 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
587 TGCCCTCCACAGCAGAGGCTGCTGTATGTGACTCTCTCAGCTTCAGAA 636
199 gLysIleIleArgGluArgGluGluGlnIleTyrTyrArgArg 213
637 GAAATCAACAGGCGCAAGGATTCTTGCATCGTGGGAACAAG 679
seq_name: gb_pl.STPSST
seq documentation_block:
LOCUS STPSST 935 bp mRNA PLN 16-MAR-1999
DEFINITION S.tuberosum mRNA for NADH-ubiquinone oxidoreductase subunit.
ACCESSION X96671
VERSION X96671.1 GI:1235606

```

```

KEYWORDS      NADH-ubiquinone oxidoreductase; PSST gene.
SOURCE        Potato.
ORGANISM      Solanum tuberosum
REFERENCE     1 (bases 1 to 935)
AUTHORS      Heiser,V., Brennicke,A. and Grolmann,L.
TITLE        The plant mitochondrial 22 kDa (PSST) subunit of respiratory chain
              complex I is encoded by a nuclear gene with enhanced transcript
              levels in flowers
JOURNAL       Plant Mol. Biol. 31 (6), 1195-1204 (1996)
MEDLINE      97071689
REFERENCE     2 (bases 1 to 935)
AUTHORS      Heiser,V.
TITLE        Direct Submission
JOURNAL       Submitted (15-MAR-1996) V. Heiser, Inst.f. Genbiologische Forschung
              Berlin, Ihnestr. 63, D-14195, Berlin, FRG
FEATURES      Location/Qualifiers
               1..935
                /organism="Solanum tuberosum"
                /cultivar="Desiree"
                /db_xref="taxon:4113"
                /clone_lib="lambda ZAP II"
                /tissue_type="green leaves"
                50..691
                 /gene="PSST"
                50..691
                 /gene="PSST"
                /EC_number="1.6.5.3"
                /note="22kDa mitochondrial complex I"
                /codon_start=1
                /product="NADH-ubiquinone oxidoreductase"
                /protein_id="CAA65451.1"
                /db_xref="GI:1235607"
                /db_xref="SWISS-PROT:Q43844"
                /translation="MALIARNKLDTAFFLORAATHITLPSLSQOPASSPATSGG
                AQPSMNTFAGISPAEYISKVDLMMWARRGSIWPMTEGLACCAVEMHGAARID
                FDRFGIIFRPSRQSDVMIVAGTLTNKMAPALRKVDMPPEPRWVISMGCANGGGY
                HYSYAVVRGCDRIVPVDIYVPGCPPTAEALLYGLLQKKINRRKDLLMWMTQ"
                50..142
                 /gene="PSST"
                 /evidence=experimental
                 143..588
                  /gene="PSST"
                  583..888
                   polyA_signal
                   BASE COUNT 221 a 212 c 221 g 281 t
                   ORIGIN
alignment_scores:
  Quality: 732.50      Length: 206
  Ratio: 4.234        Gaps: 4
  Percent Similarity: 83.981      Percent Identity: 67.961
alignment_block:
US-09-525-867-1 x STPSST
Align seg 1/1 to: STPSST from: 1 to: 935
17 LeuArgSerSerValGlyProAlaValGln...AlaArgGlyValHisG1 32
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 CTGCTCACCGGAACGGCGCGTCTCTCAAAGCGCGCGACGATCCACAC 126
32 nSerValAlaThrAspGlyProSer...SerThrGlnProAlaLeuProL 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 GACACTC.....CCATCTCTCTCACAGCAACCTGCTCTCTC 164
48 ysAlaArgAlaValAlaProLysPro..... 56
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 CGGCACCTTCAGGTGGCGACACGCTCCGTCCTCATATACACGACGAGG 214
57 ...SerSerArgGlyGluTyrValValAlaLysLeuAspLeuValAs 72

```



```

:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
601 CTCCTCCAATTGCAGAGAGAGATCAATCGGCGCAAGGACITTCCTCCTGTG 650
210 PTYRARGATG 213
|:::|
651 GTGGACAAG 660
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 17:41:45 ; Search time 1339.84 seconds
(without alignments)
10145.746 Million cell updates/sec

Title: US-09-525-867-9

Perfect score: 824

Sequence: 1 cggctgagcggctcgagcg.....cgggcaaaaaaaaaaaaaa 824

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.om.*
20: em.or.*
21: em.ov.*
22: em.pat.*
23: em.ph.*
24: em.pl.*
25: em.ro.*
26: em.sts.*
27: em.sy.*
28: em.un.*
29: em.vi.*
30: em.htgo_hum.*
31: em.htgo_inv.*
32: em.htgo_rod.*
33: em.htg_hum.*
34: em.htg_inv.*
35: em.htg_rod.*
36: em.htg_othr.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Result No.	Score	Query Match	Length	DB	ID	Description
1	736	89.3	794	9	BC001715	BC001715 Homo sapi
2	700	85.0	787	9	BC005954	BC005954 Homo sapi
3	366	44.4	754	9	AF060512	AF060512 Homo sapi
4	190	23.1	34875	9	AC005329	AC005329 Homo sapi
5	38	4.6	752	4	BTPSSTSU	X65020 B.taurus mt
6	38	4.6	329709	1	AF002997	AF002997 Mesorhizo
7	35	4.2	226013	2	AC073680	AC073680 Mus muscu
8	28	3.4	68808	2	AC022708	AC022708 Homo sapi
9	28	3.4	172631	2	AC012566	AC012566 Homo sapi
10	28	3.4	198491	2	AC022249	AC022249 Homo sapi
11	27	3.3	789	6	AX106251	AX106251 Sequence
12	27	3.3	789	6	AX140542	AX140542 Sequence
13	27	3.3	1038	6	AX008551	AX008551 Sequence
14	27	3.3	1038	9	HOMLBP	L48925 Homo sapien
15	27	3.3	1879	10	AF068261	AF068261 Rattus no
16	27	3.3	3551	10	D88190	D88190 Rattus norv
17	27	3.3	158983	2	AC020598	AC020598 Homo sapi
18	27	3.3	179206	9	CNS01DS6	AL121656 BAC seque
19	27	3.3	200368	9	CNS01DUH	AL133244 BAC seque
20	26	3.2	897	8	AF281035	AF281035 Lupinus l
21	26	3.2	911	8	BONADHPST	X82274 B.cleracea
22	26	3.2	3739	1	PDENOUVR	M93015 Paracoccus
23	26	3.2	19614	1	AF029365	AF029365 Rhodobact
24	26	3.2	138705	2	AC068923	AC068923 Oryza sat
25	26	3.2	14201	2	AC091774	AC091774 Oryza sat
26	26	3.2	153673	2	AF003769	AF003769 Oryza sat
27	25	3.0	2924	8	AF279902	AF279902 Lupinus l
28	25	3.0	144593	8	AC021891	AC021891 Genomic S
29	25	3.0	164192	9	AC074338	AC074338 Human Chr
30	25	3.0	179488	9	AC083863	AC083863 Homo sapi
31	24	2.9	224	8	AY022669	AY022669 Oryza sat
32	24	2.9	224	8	AY022809	AY022809 Oryza sat
33	24	2.9	798	9	HSA250562	AJ250562 Homo sapi
34	24	2.9	1792	9	HUMMRNAB	L10373 Human (clon
35	24	2.9	1829	9	AB062057	AB062057 Homo sapi
36	24	2.9	6043	9	AB037720	AB037720 Homo sapi
37	24	2.9	10311	1	AF008339	AF008339 Agrobacte
38	24	2.9	10898	1	AF005781	AF005781 Caulobacte
39	24	2.9	116275	2	AF003993	AF003993 Oryza sat
40	24	2.9	122215	2	AF003811	AF003811 Oryza sat
41	24	2.9	139487	2	AF003513	AF003513 Oryza sat
42	24	2.9	141017	8	AC037197	AC037197 Oryza sat
43	24	2.9	142101	2	AC091811	AC091811 Oryza sat
44	24	2.9	144798	8	AC026758	AC026758 Oryza sat
45	24	2.9	178024	8	AC025783	AC025783 Oryza sat

ALIGNMENTS

RESULT 1
LOCUS BC001715 794 bp mRNA PRI 12-JUL-2001
DEFINITION Homo sapiens, Similar to CG9172 gene product, clone MGC:886
IMAGE:3503064, mRNA, complete cds.
ACCESSION BC001715
VERSION BC001715.1 GI:12804588
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 794)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

COMMENT

Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandon, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: b Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7688970.

FEATURES

source

1..794
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:886 IMAGE:3503064"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="Vector: pORF7"
6..647
/cdon_start=1
/product="Similar to CG9172 gene product"
/protein_id="AA01715.1"
/db_xref="GI:12804589"
/translation="MAYLSAPGLRGLRSLSSVGPVAVQGVHQSVATDGPSSSTP
ALPKAVAPKPSRGEIVYAKLDLVNWARSSLWMTFLGACVEMHMAAPRYD
MORFVYFASRQSDIMVAGLTNKPALRYVDQPEPVYVSMGSCANGGYY
HYSYVRCSDRIVPDVIFPGPTTALYLIQIKRRLRIQIYWR"

BASE COUNT
ORIGIN

Query Match 89.3%; Score 736; DB 9; Length 794;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 ccaagatggcgggtgctgacgtcctggcctggcggcttcggatccttgctgcgct 133
Db 1 CCAGATGGCGGGTGTGTACGTCTCCGGCTGCGCGGCTCCGGATCTTGTCTGCCGT 60
Qy 134 ccagcgtggccggctgtgacgacgacgaggtgtccatcagagcgtggccaccgagtgcc 193
Db 61 CCAGCGTGGCGCGGCTGTGCAGGCACGAGGTGTCCATCAGAGCGTGGCCACCGATGCC 120
Qy 194 caagcacaccagcgcctgcctccaaagccagacccgtgctccaaacccagacc 253
Db 121 CAACACACCCAGCCGCGCTCCAAAGCCAGAGCCGCTGCTCCAAACCCAGACGCC 180
Qy 254 gggcgagtagtggtggccaaagtggatgacctctgaactcgaactggccggcggagttctc 313
Db 181 GGGCGAGTATGTGGTGGCCAGCTGGATGACCTGCTCAACTGGCGCGCGGAGTTCTC 240
Qy 314 tgtggccatgaacttgctgctgctgctgctgctgctgctgctgctgctgctgctgct 373
Db 241 TGTGGCCCATGACCTTGGCGCTGGCTGCTGGCGCTGGAGATGATGACATGGCAGCAC 300
Qy 374 ccgctacgacatggacgctgttggcggtgttcttcgcgcgcgcgcgcgcgcgcgcgcgcg 433
Db 301 CCCGCTAGACATGGACCGGCTTTGGGTGCTCTCCGCGCCAGCCGCGCCAGTCCGACG 360
Qy 434 tcatgactgtggcgccgacactcaccacaaagatggccccagcgttcgaaagtcctacg 493

Db 361 TCATGATGTCGGCCGCACACTCACCACAAGATGCCCCACGCTTCGCAAGGTCTACG 420
Qy 494 accagatggcggagcgcgctacgtgtgtctccatggggagctgcgccacagagagct 553
Db 421 ACCAGATGCGGAGCGCGCTACGTGTCTCCATGGGAGCTGCCCAACGAGGAGCT 480
Qy 554 actaccactattctactcgtggtgaggggtgaggggtgaggggtgaggggtgaggggt 613
Db 481 ACTACCCTATTCTTCTACTCGGTGTGAGGGGTGCGACCGCATCGTCCGCTGGACATCT 540
Qy 614 acatccagagctgccacactacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 673
Db 541 ACATCCAGGCTGCCACCTACGCGCGAGCGCGCTGTCTACGGCATCTCTGACAGTGCAGA 600
Qy 674 ggaagatcaaggggagggagggagggagggagggagggagggagggagggagggaggg 733
Db 601 GGAAGATCAAGCGGGAGCGAGGCTGCAGATCTGTGTACCGAGGTAGCGCGCGCGCGCC 660
Qy 734 gccgcggagcctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 793
Db 661 CGCGCGGAGCGCTGTGCGGTCTGTGCCAGCGCTGTGTGTCGCGTGGAGGTGTGTCAAT 720
Qy 794 aaactcgccctcgccg 809
Db 721 AAACCTGCCCTCGGGC 736

RESULT 2

BC005954
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdgaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 21 Row: k Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..787
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:14592 IMAGE:4276489"
/tissue_type="Brain, primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"

FEATURES
source

```
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS Mao, Y.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
source
1..754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..554
/codon_start=1
/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="GI:12001974"
/translaton="MAVLGAPGLRGFRILGLRSSVGLAVQARGVHQSVDGSPSSTOP
ALPKARAVAPKPSRGEYVAKLDLVNWRSSLPWMTFGLACCAVEMHMAAPRYD
MDRFGVVFRASPRQSDVMIVAGTLTKMKAPALRKSPRPDAGAAALGLHGLRQGRRL
LPFLJGEGELRPHRHLPRLLPTALLYGILOLQKIRRRRLQIWR"
BASE COUNT 152 a 262 c 244 g 129 t
ORIGIN
Query Match 85.0%; Score 700; DB 9; Length 787;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 59 gtctgaagccgagcccaagatggcgtgtgtctgaagctctgcctgcgcgggtcttcgga 118
Db 1 gtctgaagccgagcccaagatggcgtgtgtctgaagctctgcctgcgcgggtcttcgga 60
QY 119 tcttggtctgcctcagcgtggcgcgcgtgtgcagcagcaggtgtccatcagagcg 178
Db 61 tcttggtctgcctcagcgtggcgcgcgtgtgcagcagcaggtgtccatcagagcg 120
QY 179 tggccaccgcatggcccaagcagcaccagcctgcctgcgcgaagcgcagagccgtggc 238
Db 121 tggccaccgcatggcccaagcagcaccagcctgcctgcgcgaagcgcagagccgtggc 180
QY 239 ccaaacccagcagcggcgagctatgtgtggaagcctggtgacacctgtaactggg 298
Db 181 ccaaacccagcagcggcgagctatgtgtggaagcctggtgacacctgtaactggg 240
QY 299 cccgcggagttctctgtggcccatgacctggcctgcgcctgcgcgcgtggagatga 358
Db 241 cccgcggagttctctgtggcccatgacctggcctgcgcctgcgcgcgtggagatga 300
QY 359 tgcacatggcagcagcccgctacacatgagcgcgtttggcgtgttcttcgcgcgcagcc 418
Db 301 tgcacatggcagcagcccgctacacatgagcgcgtttggcgtgttcttcgcgcgcagcc 360
QY 419 cgcgcagtcgacgtcatgatcgtgcgcgcgcacactcaccacaagatggcccaagcgc 478
Db 361 cgcgcagtcgacgtcatgatcgtgcgcgcgcacactcaccacaagatggcccaagcgc 420
QY 479 ttgcgaagttctacgaccagatgcgcgagcgcgctacgtgtgtctccatgggagctgg 538
Db 421 ttgcgaagttctacgaccagatgcgcgagcgcgctacgtgtgtctccatgggagctgg 480
QY 539 ccaacggaggaggtactacacatttctactcgtgtgtgaggggctgcgacccagctcg 598
Db 481 ccaacggaggaggtactacacatttctactcgtgtgtgaggggctgcgacccagctcg 540
QY 599 tgcgcgtggacatctacatccagcgtgcacactacgcgcgagcgcctctctacggca 658
Db 541 tgcgcgtggacatctacatccagcgtgcacactacgcgcgagcgcctctctacggca 600
QY 659 tctcagctcagaggaagatcaagcgggagggaggtgcagatctggtaccgcaggt 718
Db 601 tctcagctcagaggaagatcaagcgggagggaggtgcagatctggtaccgcaggt 660
QY 719 agcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 778
Db 661 agcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
QY 779 cgtgaggtgtgcaataaaacctgcccctgggc 809
Db 721 cgtgaggtgtgcaataaaacctgcccctgggc 751
```

RESULT 3
AF060512

RESULT 4

AC005329
LOCUS AC005329 34875 bp DNA PRI 28-JUL-1998
DEFINITION Homo sapiens chromosome 19, cosmid R34382, complete sequence.
ACCESSION AC005329

```
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS Mao, Y.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
Location/Qualifiers
1..754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..554
/codon_start=1
/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="GI:12001974"
/translaton="MAVLGAPGLRGFRILGLRSSVGLAVQARGVHQSVDGSPSSTOP
ALPKARAVAPKPSRGEYVAKLDLVNWRSSLPWMTFGLACCAVEMHMAAPRYD
MDRFGVVFRASPRQSDVMIVAGTLTKMKAPALRKSPRPDAGAAALGLHGLRQGRRL
LPFLJGEGELRPHRHLPRLLPTALLYGILOLQKIRRRRLQIWR"
BASE COUNT 120 a 267 c 239 g 128 t
ORIGIN
Query Match 44.4%; Score 366; DB 9; Length 754;
Best Local Similarity 99.8%; Pred. No. 2e-171;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 70 gagcccaagatggcgggtgtctcagctcctggcctgcgcgcgttcggatccttggtctg 129
Db 1 GAGCCAAAGATGGCGGTGTCTCAGCTCCTGGCCTGCGCGCTCCGATCCTTGTTCTG 60
QY 130 cgtctcagctgggcccggctgtcaggcagcagaggtgtccatcagagcgtggccacgat 189
Db 61 CGCTCCACGCTGGCGCTGTGCTGCAGCAGAGGTCTCCATCAGAGCGTGGCCACCGAT 120
QY 190 ggcccaagcagcaccagcctgcctgcacaaaggccagcgcgtggctcccaaccagc 249
Db 121 GGCCCAAGCAGCACCCAGCCTGCCCTGCCAAAGCCAGACCGCTGCCCAACCCAGC 180
QY 250 agccggggcagatgtgtgtggccaagctggatgacctgtcaactgggcccgcggagt 309
Db 181 AGCGGGGCGAGTATGTGTGGCCAAAGTGGATGACCTCGTCAACTGGCGCCCGGAGT 240
QY 310 tctctgtggcccatgacctgcgcgtgcctgcctgcctgcctgcctgcctgcctgcctgc 369
Db 241 TCTCTGTGGCCCATGACCTTCGGCTGCGCTGCTGCGCCGTGGAGATGATGCACATGGCA 300
QY 370 gcacccgcctcagcagatggacccgttttggcgtgttcttcgcgcgcagccgcgcagtc 429
Db 301 GCACCCGCTACGACATGACCGCTTTTGGCGTGTCTTCGCGCCAGCCGCGCCAGTCC 360
QY 430 gacgtcagtcgtggccggcacactcaccaaaagatggccccagcgttccag 486
Db 361 GACGTCATGATCGTGGCGGCACACTCACCACAAAGATGCCCCCAGCGCTTCGCAAG 417
```

```

VERSION      AC005329.1  GI:3342732
KEYWORDS     HFG.
SOURCE       Human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 34875)
              Lamerdin, J.E., McCready, P.M., Skowronski, E., Adamson, A.W.,
              Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stillwagen, S.,
              Phan, H., Velasco, N., Do, L., Regala, N., Terry, A., Garnes, J.,
              Dangnanan, L., Poundstone, P., Christensen, M., Georgescu, A., Avila, J.,
              Liu, S., Attix, C., Andrejsek, T., Frankheim, M., Amico-Keller, G.,
              Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G.,
              Krommiller, B., Arellano, A., Montgomery, M., Ow, D., Nolan, M.,
              Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V.
TITLE        Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
              serine protease gene cluster
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 34875)
AUTHORS      Lamerdin, J.E.
TITLE        Direct Submission
JOURNAL      Submitted (28-JUL-1998) Joint Genome Institute, Lawrence Livermore
              National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
COMMENT      Map and sequence oriented from p telomere to
              centromere. Cosmid R34382 is adjacent to cosmid R28991 to the left,
              and overlaps BAC 38854 to the right from bases 33,973 to 34,873.
              Estimated sequence gap between R34382 and R28991 is <1 kb.
              Additional map and sequence information may be obtained at:
              http://www-bio.lnl.gov/bbrp/genome/genome.html.
FEATURES     Location/Qualifiers
              1..34875
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="R34382"
                /chromosomes="19"
                /map="19p13.3 between CDC34 and D19S342"
                /cell_line="5HL2-B"
                /clone_lib="LL19NC03 R chromosome 19-specific cosmid
                library"
                /note="Cosmid library constructed at LLNL from flow-sorted
                chromosomes from human-hamster hybrid 5HL2-B, which
                carries chromosome 19 as its only human chromosome."
              repeat_region complement(1..62)
                /rpt_family="Alu"
              misc_feature 117..921
                /note="DDS similarity to overlapping ESTs:
                (564..117) AA128278 zw18a09.s1 Soares ovary tumor NBshot
                Homo sapiens cDNA clone 769524 3'. Score: 843 Identity:
                441/461 (95%).-(563..127) AA573130 nj41f11.s1 NCI_CGAP_AA1
                Homo sapiens cDNA clone IMAGE:995085. Score: 641
                Identity: 441/446 (98%).-(541..99) H89953 ys82h05.s1 Homo
                sapiens cDNA clone 22137 3'. Score: 733 Identity:
                422/464 (90%).-(148..623) N31319 yx34a05.r1 Homo sapiens
                cDNA clone 265520 5'. Score: 930 Identity: 474/478 (99%).
                -(387..744) C117502 Human placenta cDNA 5'-end GEN-548D09.
                Score: 690 Identity: 354/356 (99%).-(458..921) R14335
                yf79p09.r1 Homo sapiens cDNA clone 28433 5'. Score: 592
                Identity: 440/496 (88%).-Additional EST matches:
                D44938, H71984, D81984, R35196"
              misc_feature 128..181
                /note="Predicted exon, program: graill2exons_human_1.3,
                frame: 2, quality: good, score: 72.000"
              misc_feature complement(1518..2040)
                /note="DDS similarity to H98977 yx11e03.s1 Homo sapiens
                cDNA clone 261436 3'. Score: 967 Identity: 509/522
                (97%).-Additional EST matches:
                AA335644, AA677688, R50856, AA969114, R40661, T09489"
              repeat_region 2201..2486
                /rpt_family="AluSx"
              repeat_region 2516..2627
                /rpt_family="AluJb"
              repeat_region complement(3800..3931)
                /rpt_family="(CCAA)n"
              repeat_region complement(3832..4090)
                /rpt_family="AluJb"
              repeat_region complement(4107..4200)
                /rpt_family="(GGGA)n"
              repeat_region complement(4398..4553)
                /rpt_family="MLT1B"
              repeat_region complement(4617..4667)
                /rpt_family="MLT1B"
              repeat_region 4670..4967
                /rpt_family="AluSx"
              repeat_region 5229..5500
                /rpt_family="T1MC3"
              repeat_region complement(5527..5559)
                /rpt_family="AT-rich"
              repeat_region 5572..5866
                /rpt_family="AluSx"
              repeat_region complement(5983..6109)
                /rpt_family="FLAM_C"
              repeat_region complement(6294..6551)
                /rpt_family="AluSg/X"
              repeat_region complement(6971..7043)
                /rpt_family="tRNA-Phe-TTC"
              repeat_region 7172..7246
                /rpt_family="tRNA-Asn-AAC"
              repeat_region 7536..7573
                /note="Predicted exon, program: graill2exons_human_1.3,
                frame: 2, quality: excellent, score: 79.000"
              repeat_region 7558..7657
                /rpt_family="(CGG)n"
              repeat_region 8050..8421
                /rpt_family="MER80"
              repeat_region complement(8422..8696)
                /rpt_family="AluJb"
              repeat_region 8848..9142
                /rpt_family="AluSp"
              repeat_region 9147..9447
                /rpt_family="AluSp"
              repeat_region complement(9856..10091)
                /rpt_family="MLT1B"
              repeat_region complement(10109..10403)
                /rpt_family="AluSx"
              repeat_region complement(10407..10508)
                /rpt_family="MLT1B"
              repeat_region complement(12031..12535)
                /note="DDS similarity to AA594148 nn3le07.s1 NCI_CGAP_Gas1
                Homo sapiens cDNA clone IMAGE:1085508 similar to
                SW:NUK_BOVIN P42026 NADH-UBIQUINONE OXIDOREDUCTASE 20 KD
                SUBUNIT PRECURSOR ; Score: 960 Identity: 495/503 (98%)."
              CDS join(<12485..12547,14480..14659,14728..14774,16851..16939,
              19000..19097)
                /note="NADH-UBIQUINONE OXIDOREDUCTASE 20 KD SUBUNIT
                PRECURSOR; COMPLE: C1-20KD; PSST SUBUNIT"
                /codon_start=1
                /product="NUK_HUMAN, partial CDS"
                /protein_id="AAC27669.1"
                /db_xref="GI:3342734"
                /translation="PSRGENYVAKLDLDLVNWRSSLPWMTCLACAVEMHMRAP
                RYDMRFQVFRPRSDMIVAGTTLNMAPLRKVYDQMPPEPRVYVSGSCANGG
                GYHYYSVVRGCDRIVPVDIYIPGCPPTAALLYILQIRKRRRIQIYYRR"
              repeat_region 12596..12649
                /rpt_family="(CA)n"
              repeat_region 12653..12784
                /rpt_family="(CATG)n"
              repeat_region 12785..12872
                /rpt_family="(CA)n"
              repeat_region 12873..13006
                /rpt_family="(CATG)n"
              repeat_region complement(13512..13807)
                /rpt_family="AluY"
              repeat_region 14197..14296
                /rpt_family="MER33"
              repeat_region complement(15093..15391)
                /rpt_family="AluSx"

```


AP002997 BA000012
 AP002997.2 GI:14022051
 Mesorhizobium loti (strain:MAFF303099) DNA.
 Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 Phyllobacteriaceae; Mesorhizobium.
 1 (sites)
 Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,
 Watanabe,A., Idesawa,K., Ishikawa,A., Kawashima,K., Kimura,T.,
 Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,
 Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpo,S., Sugimoto,M.,
 Takeuchi,C., Yamada,M. and Tabata,S.
 Complete genome structure of the nitrogen-fixing symbiotic
 bacterium Mesorhizobium loti
 DNA Res. 7 (5), 331-338 (2000)
 21082930
 2 (bases 1 to 329709)
 Kaneko,T.
 Direct Submission
 Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research, Yana
 1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail:kaneko@kazusa.or.jp/rhizobase/)
 URL: <http://www.kazusa.or.jp/rhizobase/>
 Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
 On May 11, 2001 this sequence version replaced gi:11994965.
 Location/Qualifiers
 1..329709
 /organism="Mesorhizobium loti"
 /strain="MAFF303099"
 /db_xref="taxon:381"
 complement(79..402)
 /gene="mlr1244"
 complement(79..402)
 /note="unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB48662.1"
 /db_xref="GI:14022052"
 /translation="MSRKGSLPFGNSFHYWADCLMVQRPIMSDLLLSIFTFTVMR
 LLKGLWLNPOYLATGILGAVVAMLLNGLPAYDDDFVIGGVGTGFGWAGNALFEDA
 ILGIA"
 411..2480
 /gene="mlr1245"
 411..2480
 /gene="mlr1245"
 /codon_start=1
 /transl_table=11
 /product="probable electron transfer protein"
 /protein_id="BAB48663.1"
 /db_xref="GI:14022053"
 /translation="MNSPANTIDPLVLFMPGSKGRFPVGTPTVLDARQLGVYVESVC
 LGRDVCQGEVEQNEFPAKHIVSSNDHISPGPKKEERYVERVGPERRRSCSAQI
 LGDLVFPQDTVINAQTRKADATVRIADTAIRMVCYVEIEEPDMHKPLGLDRLKI
 ALMDQWGFKEFEFFYLLPOVQGITLRGNWTATAAIHKDADSDIARIALWPLGKNEA
 YGLACDGSITIAHLVSLSGRVAASGTSNPQIRGEDLMKRSVYVMNPDGREGM
 TVAVEATSSIVDKVCAQGNVQRNIDLSVFVGNPIMHFLGLDPTELGGAFFALAV
 SGAVIKASDGLKINOCARLYMLPCITAGHVGADAAVTISEGPHQDQEMLLIVDGT
 NAEIVLNRARVVAASSTGPAFGAIEISGQRAAGIAERIDPTELEPKYRIGS
 ELWSDPFLDSVQATGVTGICGSGIIEIIVAEYLAGIISDGVGDSLAARSPRVTA
 NGRFSYVKEGEKITITQDVRAIQAKAALYAGTKLLMEKQNTHEVDRIHFAGAF
 GSFIDPKYAMVLGILPCDDLDKVSAGVNAAGAGAMALLNRYRREITEETVSOIEKIE
 TALEPKGHEFVYAMALPNKVDPPKLSAAVKLPKRTVSEDGIAGDATPRRSREGH
 AARRSRE"
 2529..2852
 /gene="mlr1246"
 2529..2852
 /gene="mlr1246"
 /note="unknown protein"
 /codon_start=1
 /transl_table=11

/protein_id="BAB48664.1"
 /db_xref="GI:14022054"
 /translation="MFPDRTDPTETTFSLRNLSDTQWHLPRIERPERPEILRRSSRE
 MKMSFENLGRFGVTIRQHARNKAIKALNSLFAHVQKIDGWSPSPSPDQAALASLL
 LGTAR"
 3087..4040
 /gene="mlr1247"
 3087..4040
 /gene="mlr1247"
 /codon_start=1
 /transl_table=11
 /product="esterase"
 /protein_id="BAB48665.1"
 /db_xref="GI:14022055"
 /translation="MPSIKSHFVSFLVRSKKAFFSPENLRBWLIAARKTDDHHPA
 SLHGRHLIQTRTVGDFPFVYAPKAGQRRILHLHGAYVFEITPPHHWLIADMBRL
 GYGITVPIYIAPSHDFHAFMGVMDYROMLDETAEDIVFEMFGDGGHNNVLLTMM
 AAEGLPLPAHVLISPLGMSLNPKLFAERNDPWLGIAGGLEAIRMYSGIDRSR
 WHISPIYGDVSLPTLLTGTSHDLTPDNLIFAQKARDAGVEVEVVEEGMFWHAPL
 IDMEARRARQHIVAFILGENRSPVRQDGKTRFOAPSAAAE"
 complement(4062..4961)
 /gene="mlr1248"
 complement(4062..4961)
 /gene="mlr1248"
 /codon_start=1
 /transl_table=11
 /product="probable transcriptional regulator"
 /protein_id="BAB48666.1"
 /db_xref="GI:14022056"
 /translation="MNAPLNHPLLDLVLRFTVAETGSETTAANAVERTPSAVS
 MOIKLEDLGORSVPARDARSVSLTGDGMLGYARALLSINREVSVKFIPIPVGV
 RLGSFDYDGERVLPVLRFAQSHPSIADVTIDQSNLRMRDDRALDITLLINYSK
 TSALGAEVLLTPEIVWAGAKGCAHLREPLFVSLWEEGCARAGALEALGREGNRYV
 AYMSAHTAGQRAIMADLAVAPLPRSEFLGNDMVLCFGKMGPDIGTYNLAMVYAPAS
 APKAVADHIRATEFVFRETGKF"
 5113..5379
 /gene="msr1249"
 5113..5379
 /gene="msr1249"
 /note="unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB48667.1"
 /db_xref="GI:14022057"
 /translation="MGLKHAACKTSQIKRNVPGPAPKENMSILSSLGRIATEFSA
 RARYQTBRAIRSLFIELOKIDGWPEASDTGTGRNGVSWAGAK"
 6136..7137
 /gene="mlr1251"
 6136..7137
 /gene="mlr1251"
 /codon_start=1
 /transl_table=11
 /product="probable transcriptional regulator"
 /protein_id="BAB48668.1"
 /db_xref="GI:14022058"
 /translation="MNAIKHPIKRPSFVFLVPDFTMFATALDPLRSANRMUGYEA
 FWRLLASIDGKVRASNGECAVNTSLPEERKKNAGDRPNMAIVCGSVNRYQNKSA
 FAWLRVENRVAVGGCTGAHILAAAGLSNKRCAIHMENLPGFSEAPKANVPADL
 FEVDQNYTCAGGTAALDMKLKIGDDDFDSLNVRYCEQVLTDRVSPDROHPLURA
 RLGVQNSKVTIIEELMEGNLSEPLSLIEIADHDVLSRRQJERLFRTEMGRSPARYLE
 IRLDARHLLIQSSMPVVEVAVACGVSASHFSKCIKRELYARSFQQRVDRKQLAA"
 complement(7153..7956)
 /gene="mlr1253"
 complement(7153..7956)
 /gene="mlr1253"
 /codon_start=1
 /transl_table=11
 /product="ATP-binding protein of ABC transporter"
 /protein_id="BAB48669.1"
 /db_xref="GI:14022059"
 /translation="MEPCVSKLLIEGVSRITFAGVYRGLPVKALMPIDLVAANDFTIT
 LGPSGCKSTLLRIVAGLEAPSEGRVLLDDKAVTRPGPDGRGMVFOSYTLFPWLTUSE
 IAFGLRGRMAPKERDDIVASYVDLVGLKGFENHPKQLSGGMQQTALARALANDE


```

FEATURES
  source
    * 70375 226013: contig of 155639 bp in length.
      Location/Qualifiers
        1..226013
          /organism="Mus musculus"
          /db_xref="taxon:10090"
          /clone_lib="RP23-116Mi"
          /clone_lib="RPCI mouse BAC library 23"
BASE COUNT 54098 a 60347 c 59323 g 50104 t 1941 others
ORIGIN

Query Match
Best Local Similarity 4.2%; Score 35; DB 2; Length 226013;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 cagaggttctcgaagccgagccgaagtcgag 85
|||||
Db 182933 CAGAGGTTGTCGAGCCGAGCCGAGGATGCGG 182799

RESULT 8
AC022708 68808 bp DNA HTG 24-MAR-2001
LOCUS Homo sapiens clone RP11-76H7, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC022708
ACCESSION AC022708
VERSION AC022708.3 GI:13443223
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 68808)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome, clone RP11-76H7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 68808)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArrelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lied,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 24, 2001 this sequence version replaced gi:9146492.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5354
Center clone name: 76_H7
-----
* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely

```

```

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 717: contig of 717 bp in length
* 718 817: gap of 100 bp
* 818 1550: contig of 733 bp in length
* 1551 1650: gap of 100 bp
* 1651 2358: contig of 708 bp in length
* 2359 2458: gap of 100 bp
* 2459 3179: contig of 721 bp in length
* 3180 3279: gap of 100 bp
* 3280 4009: contig of 730 bp in length
* 4010 4109: gap of 100 bp
* 4110 4816: contig of 707 bp in length
* 4817 4916: gap of 100 bp
* 4917 5635: contig of 719 bp in length
* 5636 5735: gap of 100 bp
* 5736 6456: contig of 721 bp in length
* 6457 6556: gap of 100 bp
* 6557 7286: contig of 730 bp in length
* 7287 7386: gap of 100 bp
* 7387 8118: contig of 732 bp in length
* 8119 8218: gap of 100 bp
* 8219 8933: contig of 715 bp in length
* 8934 9033: gap of 100 bp
* 9034 9769: contig of 736 bp in length
* 9770 9869: gap of 100 bp
* 9870 10593: contig of 724 bp in length
* 10594 10693: gap of 100 bp
* 10694 11426: contig of 733 bp in length
* 11427 11526: gap of 100 bp
* 11527 12238: contig of 712 bp in length
* 12239 12338: gap of 100 bp
* 12339 13050: contig of 712 bp in length
* 13051 13150: gap of 100 bp
* 13151 13868: contig of 718 bp in length
* 13869 13968: gap of 100 bp
* 13969 14699: contig of 731 bp in length
* 14700 14799: gap of 100 bp
* 14800 15512: contig of 713 bp in length
* 15513 15612: gap of 100 bp
* 15613 16336: contig of 724 bp in length
* 16337 16436: gap of 100 bp
* 16437 17180: contig of 744 bp in length
* 17181 17380: gap of 100 bp
* 17381 17986: contig of 706 bp in length
* 17987 18086: gap of 100 bp
* 18087 18800: contig of 714 bp in length
* 18801 18900: gap of 100 bp
* 18901 19628: contig of 728 bp in length
* 19629 19728: gap of 100 bp
* 19729 20442: contig of 714 bp in length
* 20443 20542: gap of 100 bp
* 20543 21248: contig of 706 bp in length
* 21249 21348: gap of 100 bp
* 21349 22075: contig of 727 bp in length
* 22076 22175: gap of 100 bp
* 22176 22898: contig of 723 bp in length
* 22899 22998: gap of 100 bp
* 22999 23723: contig of 725 bp in length
* 23724 23823: gap of 100 bp
* 23824 24543: contig of 720 bp in length
* 24544 24643: gap of 100 bp
* 24644 25430: contig of 787 bp in length
* 25431 25530: gap of 100 bp
* 25531 26351: contig of 721 bp in length
* 26352 26351: gap of 100 bp
* 26352 27071: contig of 720 bp in length
* 27072 27171: gap of 100 bp

```


* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 78380: contig of 78380 bp in length
 * 78381 78480: gap of 100 bp
 * 78481 172631: contig of 94151 bp in length.

FEATURES

source
 1. 172631
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP11-38D12"
 /clone_lib="RPC1-11 Human Male BAC"
 1. 78380
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 78481..172631
 /note="assembly_fragment
 clone_end:17
 vector_side:right"
 52774 a 30937 c 31576 g 57244 t 100 others

BASE COUNT

ORIGIN

Query Match 3.4%; Score 28; DB 2; Length 172631;
 Best Local Similarity 100.0%; Fred. No. 0.0024;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 797 cctgcctcggcgcaaaaaa824

Db 91239 CCGCCCTCGGCAAAAAA91212

RESULT 10

AC022249
 LOCUS
 DEFINITION Homo sapiens clone RP11-24614, WORKING DRAFT SEQUENCE, 19 unordered
 pieces.
 AC022249 198491 bp DNA HTG 09-SEP-2000
 AC022249.3 GI:10045487
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 198491)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Aldrich,E., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckler,R., Beda,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
 Chael,Y., Colangelo,M., Collins,S., Collins,A., Cooke,P.,
 DeArnellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lehotzky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 McPheters,R., Meldrim,J., Menus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
 Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye.W.J.,
 Zimmer,A. and Zody,M.

TITLE JOURNAL COMMENT

Direct Submission
 Submitted (27-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 9, 2000 this sequence version replaced gi:5850453.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4451
 Center Clone name: 24_G_14
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 187097 bases at least Q40
 Consensus quality: 192525 bases at least Q30
 Consensus quality: 194606 bases at least Q20
 Insert size: 196000; agarose-fp
 Insert size: 196691; sum-of-contigs
 Quality coverage: 4.4 in Q20 bases; agarose-fp
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 19 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 23961: contig of 23961 bp in length
 23962 24061: gap of 100 bp
 24062 50896: contig of 26835 bp in length
 50897 50996: gap of 100 bp
 50997 52300: contig of 1304 bp in length
 52301 52400: gap of 100 bp
 52401 54995: contig of 2595 bp in length
 54996 55095: gap of 100 bp
 55096 57603: contig of 2508 bp in length
 57604 57703: gap of 100 bp
 57704 62215: contig of 4512 bp in length
 62216 62315: gap of 100 bp
 62316 66599: contig of 4284 bp in length
 66600 66699: gap of 100 bp
 66700 72051: contig of 5352 bp in length
 72052 72151: gap of 100 bp
 72152 79334: contig of 7183 bp in length
 79335 79434: gap of 100 bp
 79435 87876: contig of 8442 bp in length
 87877 87976: gap of 100 bp
 87977 95869: contig of 7893 bp in length
 95870 95969: gap of 100 bp
 95970 105886: contig of 917 bp in length
 105887 105986: gap of 100 bp
 105987 116314: contig of 10328 bp in length
 116315 116414: gap of 100 bp
 116415 127405: contig of 10991 bp in length
 127406 127505: gap of 100 bp
 127506 140893: contig of 1388 bp in length
 140894 140993: gap of 100 bp
 140994 159375: contig of 18382 bp in length
 159376 159475: gap of 100 bp
 159476 178521: contig of 19046 bp in length
 178522 178621: gap of 100 bp
 178622 197646: contig of 19025 bp in length
 197647 197746: gap of 100 bp
 197747 198491: contig of 745 bp in length.

Location/Qualifiers

FEATURES

TITLE	Compositions and methods for therapy and diagnosis of prostate cancer									
JOURNAL	Patent: WO 0125272-A 32 12-APR-2001; CORIXA CORPORATION (US)									
FEATURES	Location/Qualifiers 1..789 /organism="Homo sapiens" /db_xref="taxon:9606"									
misc_feature	1..789	/note="n = A,T,C or G"								195 t
BASE COUNT	129 a	240 c	159 g	195 t	66 others					
ORIGIN										
Query Match 3.3%; Score 27; DB 6; Length 789;										
Best Local Similarity 100.0%; Pred. No. 0.021; 0; Gaps 0;										
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
Qy	722	gccgcgcgcgcgcgcgcgcgcgcctgt	748							
Db	151	gccgcgcgcgcgcgcgcgcgcgcctgt	125							
RESULT 12										
AX140542/C										
LOCUS	AX140542	789 bp	DNA	PAT	31-MAY-2001					
DEFINITION	Sequence 32 from Patent WO0134802.									
ACCESSION	AX140542									
VERSION	AX140542.1	GI:14280660								
KEYWORDS	human.									
SOURCE	Homo sapiens									
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (bases 1 to 789)									
AUTHORS	Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H., Skeiky,Y.A. and Wang,A.									
TITLE	Compositions and methods for the therapy and diagnosis of prostate cancer									
JOURNAL	Patent: WO 0134802-A 32 17-MAY-2001; CORIXA CORPORATION (US)									
FEATURES	Location/Qualifiers 1..789 /organism="Homo sapiens" /db_xref="taxon:9606"									
BASE COUNT	129 a	240 c	159 g	195 t	66 others					
ORIGIN										

```
Query Match          3.3%; Score 27; DB 6; Length 789;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 27; Conservative 0; Mismatches 0; Gaps 0;
```

RESULT	13
AX008551	
LOCUS	1038 bp DNA
DEFINITION	Sequence 9 from Patent WO9966059.
ACCESSION	AX008551
VERSION	AX008551.1 GI:9996101
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; catarrhini; Hominoidea; Homo.
AUTHORS	Del-Favero, J. and Van Broeckhoven, C.
TITLE	Novel fragmentation vectors and uses thereof

Search completed: February 12, 2002, 19:09:51
Job time: 5286 sec


```

Db 515 CCCACCTAGCCGCGAGGCGCTCTCTACGACATCCCTGACCTGACAGGAGATCAGCG 574
QY 587 ggaagcgagggtgcagatctgtgtaccagcaggtagcgccg 726
Db 575 GGAGCGGAGGTGTCAGATCTGTGTACCGCAGGTAGCGCGCG 614

RESULT 6
AW245467/c 586 bp mRNA EST 07-JAN-2000
LOCUS 2822901.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822901 3',
DEFINITION mRNA sequence.
ACCESSION AW245467
VERSION AW245467.1 GI:6588460
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: 2822901.5prime
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project
Project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross_match from University of Washington Genome Center
PRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu/polyadenylation: Based upon the
presence of a XhoI site followed by a run of 14 or more T residues
at the beginning of the sequence, this cDNA insert was
polyadenylated.
Plate: LICM10 row: J column: 22
High quality sequence stop: 474.
FEATURES
Location/Qualifiers
1..586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2822901"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using 2AP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 100 a 181 c 199 g 106 t

```

Query Match 69.9%; Score 576; DB 10; Length 586;
Best Local Similarity 100.0%; Pred. No. 1.2e-152;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 gtggctcccaaccagcagcgggcgagtgatgtgtggccaagctggatgacctgtc 291
Db 586 GTGGCTCCCAACACAGCAGCGGGCGAGTATGTGTGGCCAGCTGGATGACCTCGC 527
QY 292 aactggcgccgcgaggtctctgtgcccattgaccttgcctgctgcccgtg 351
```

```

Db 526 AACTGGGCCCCCGAGATTCTGTGTGCCCATGACCTTGGCCCTGGCCCTGCTGCGCGGTG 467
QY 352 gagatgatgacatgagcagcagcccccgtacgacatgacacccgttttggtgtttccgcg 411
Db 466 GAGATGATGACATGGCAGCACCACCCGCTAGCAGATGACCGCTTTGGCGTGTCTTCGCG 407
QY 412 gccagcccgcccgagtcgagcgtcatgattggtgcccgcacactcaccacaagatggcc 471
Db 406 CCGAGCCCGCCCGCAGTCCGAGCTGATGATCGTGCCCGGACACTCACCACAGATGGCC 347
QY 472 ccagcgttgcgaaggtctacgaccagatgccggagcgcggtacgtggtctccatgggg 531
Db 346 CCAGCGTTGCAAGGCTCTACGACCCAGATCCGAGCGCGGCTACGTGCTCCATGGGG 287
QY 532 agctgcgccaacggagaggtctactaccattctactactcgtgtgagggtcgcac 591
Db 286 AGCTGGCGCAACGAGGAGGCTACTACCACTATTCCTACTCGGTGTGAGGGGCTCGGAC 227
QY 592 cgcactgtgcccgtggacatctacatcccagggtgccacctacgcgcgagggccctgctc 651
Db 226 CGCATCGTGGCCGTGGACATCTACATCCAGGCTGCCACCTACGGCCGAGGCCCTGCTC 167
QY 652 tacggatctctgcagctgcagaggaagatcaagcgggagcggggctgcagatctggtac 711
Db 166 TACGGCATCTTCAGCTCCAGAGGAAGATCAAGCGGAGCGGAGGCTCAGATCTGTATC 107
QY 712 cgcaggtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 771
Db 106 CGCAGGTAGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 47
QY 772 tgtgtcccgtaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 807
Db 46 TGTGTCCCGTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11

RESULT 7
BF342354 804 bp mRNA EST 22-NOV-2000
LOCUS 602013145F1 NCI_CGAP_Brn54 Homo sapiens cDNA clone IMAGE:4148842
DEFINITION 5', mRNA sequence.
ACCESSION BF342354
VERSION BF342354.1 GI:11289355
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 804)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9409 row: j column: 11
High quality sequence start: 2
High quality sequence stop: 774.
FEATURES
Location/Qualifiers
1..804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4148842"
/clone_lib="NCI_CGAP_Brn54"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;
```

BASE COUNT	149 a	268 c	255 g	132 t
Site_2:	Sali; Cloned unidirectionally. Primer: Oligo dt.			
	Average insert size 1.57 kb. Constructed by Life			
	Technologies. Note: this is a NCI_CGAP Library."			

```

BASE COUNT      149 a e 268 c 255 g 132 t
ORIGIN
Query Match      69.5%  Score 573;  DB 11;  Length 804;
Best Local Similarity 100.0%;  Pred. NO. 6.6e-152;
Matches 573;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

602269415F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4357803 5',
mRNA sequence.
ACCESSION BF968344
VERSION BF968344.1 GI:12335559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 936)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9994 row: m column: 04
High quality sequence stop: 704.
Location/Qualifiers
1..936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4357803"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dr
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 192 a 283 c 318 g 143 t
ORIGIN
Query Match 65.4%; Score 539; DB 11; Length 936;
Best Local Similarity 100.0%; Pred. No. 2e-142;
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 90 gtacgtctctgctgcgggttcctcgatccttggctgctgcctccagctggccggc 149
Db 40 GTCAGCTCTGCGCTGCGGGCTTCGGATCCTTGGTGTGGCTCCAGCGTGGCGCGGC 99
QY 150 ttgtcaggcacaggtgtccatcacagcgtgacccatgcccacagcagccacccagcc 209
Db 100 TGTGCAGGCACAGAGTGTCCATCAGAGGCTGCGCCACCATGGCCACAGCACCCAGCC 159
QY 210 tgccttgcacaaagccagagccgtggctcccaaacccagcagccggggcgagtgatgtgt 269
Db 160 TGCCCTGCGCAAAAGGCGAGAGCGGTGGCTCCCAAAACCCAGCAGCGGGCGAGTGTGGT 219
QY 270 gcccagctcgatgacctcgtcaactgagccgcgcgcgaggttctctgtgcccagacctt 329
Db 220 GGCCAGCTGGATGACCTCGTCACTGGCCCGCGCGAGTCTCTGTGGCCCATGACCTT 279
QY 330 cggccttggcctgctgcgcgtggagatgacacatggcagcagcaccctcagcacatgga 389
Db 280 CGGCCCTGGCTGCTGGCGCGTGGAGATGATGCATGCGCAGCACCCCGCTACGACATGGA 339
QY 390 ccgcttggcgtggttttccgcgcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 449
Db 340 CCCTTTGGCGTGGTGTTCCTCCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 399
QY 450 cacactcacacaagaatggccccgcgcttgcgaaggctctacgaccagatgcccggagcc 509
Db 400 CACACTCACACAGATGGCCCCAGCGCTTCGCAAGTCTACGACCATGATGCGCGAGCC 459

```

```

QY 510 gcgcacagcgtgtctccatcgaggagctgcgcaacgagagaggtactaccactattccta 569
Db 460 GCGCTACGTGGTCTCCATCGGGAGAGCTGGCCARCGAGAGAGCTACTACCTATTCTTA 519
QY 570 ctcggtgtgtaggggtgcgaccgcacatcgtgccgtggacatctacatccacagctgcc 628
Db 520 CTCGGTGTGAGGGGTGCGACCGCATGTCGCCGTGGACATCTACATCCCGAGGCTGCC 578

RESULT 10
BE97893 883 bp mRNA 20-OCT-2000
LOCUS 601440565F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3925168 5',
DEFINITION mRNA sequence.
ACCESSION BE97893
VERSION BE97893.1 GI:10363813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9764 row: b column: 17
High quality sequence stop: 726.
Location/Qualifiers
1..883
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3925168"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 185 a 276 c 282 g 140 t
ORIGIN
Query Match 65.0%; Score 536; DB 11; Length 883;
Best Local Similarity 100.0%; Pred. No. 1.9e-141;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 agctcctgcctgcgcggttcctccgataccttgctgcctccagcgtggcccggtgt 152
Db 1 AGCTCTGCGCTCGCGGCTTCGGATCCTTGGTGTGGCTCCAGCGTGGCGCGGCTGT 60
QY 153 gcaggcacagtgatgtccatcacagcgtggccacogatggcccaagcagccacccagctgc 212
Db 61 GCAGGCACGAGGTGTCCATCAGAGCGTGGCCACCGATGGCCCAAGCAGCAGCCAGCTGC 120
QY 213 cctgcacaaagccagagccgtgtctcccaaacccagcagccggggcgagtgatgtgtg 272
Db 121 CTGCCCCAAGAGCCAGAGCGCTGGCTCCCAACCCAGCAGCGCGGGCGAGTGTGGTGGC 180
QY 273 caagctggatgacctcgtcaactggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 332
Db 181 CAAGCTGGATGACCTCGTCAACTGGCGCGCGCGAGGTTCTCTGTGGCCCATGACCTTCG 240
QY 333 cctgcctcctgcgcgcgtgagatgatgcacatggcagcaccgccgcgcgcgcgcgcgcgc 392

```


following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |

99 a 204 c 184 g 99 t
BASE COUNT 140 a 247 c 249 g 137 t
ORIGIN
Query Match 64.8%; Score 534; DB 11; Length 586;
Best Local Similarity 99.8%; Pred. No. 7.3e-141;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 93 agctcctggtcgcggtctccggtatccttggtctgctgctccggtggtggtcgtgt 152
DB 2 agctcctggtcgcggtctccggtatccttggtctgctgctccggtggtggtcgtgt 61
QY 153 gaggcagaggtgtccatcagagcgtggtccacgagtggtccacgagcagcctgc 212
DB 62 GCAGCAGAGGTGTCATCAGAGGTGTCACCGATGGCCACGATGGCCACGACGACGCTGC 121
QY 213 cctgccaaggccagagcgtggtctccaaacccacgagcggggtggtggtggtgc 272
DB 122 COTGCCAAAGGCGACAGCGGTGGCTCCCAAAACCAGCAGCGGGCGGAGTATGTGGTGC 181
QY 273 caagctgagatgacctgtcaactggtggtggtggtggtggtggtggtggtggtggt 332
DB 182 CAAGTGTGATGACCTGCTCACTAGGCGGCGGCGGAGTCTCTGTGGCCCATGACCTTCGG 241
QY 333 cctggtcgtgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 392
DB 242 COTGCCGTGCTGCGCGGTGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 301
QY 393 ctttgggtggttctccggtcagcggcggcggcggcggcggcggcggcggcggcggc 452
DB 302 CTTTGGCGTGTGCTTCCGCGCCAGCCCGCGCGAGTCCGAGCTCATGATGATGATGATGAT 361
QY 453 actcaccacaagatggtccacgagcgttctgcaaggtctacgacagatggtggtggtggt 512
DB 362 ACTCACCACAGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 421
QY 513 ctactgtgtctcactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 572
DB 422 CTACTGTGTCTCTCACTGGGAGCTGCGCCAGCAGGAGGAGGCTACTTACCACATTCTACT 481
QY 573 ggtgtgaggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 632
DB 482 GGTGTGAGGGGTGCGACCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 541
QY 633 tacggtcagggcgtgtctacggtggtggtggtggtggtggtggtggtggtggtggtggt 677
DB 542 TACGCGCGAGGCGCTGCTCTACGGCATCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAG 586

RESULT 13
BF206581 773 bp mRNA EST 06-NOV-2000
LOCUS 601870924F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4101725 5',
DEFINITION mRNA sequence.
ACCESSION BF206581
VERSION BF206581.1 GI:11100167
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 773)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LicM973 row: o column: 06
High quality sequence stop: 744.
FEATURES
source
1. 773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4101725"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 140 a 247 c 249 g 137 t
ORIGIN

Query Match 63.8%; Score 526; DB 11; Length 773;
Best Local Similarity 100.0%; Pred. No. 1e-138;
Matches 526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 ctgctccatcagagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 162
DB 26 CTGCGCGCTTCCGGATCCTTGGTCTCGCTCCAGCTGGCGCGCGCTGGCGACGACGA 85
QY 163 ggtgtccatcagagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 222
DB 86 GGTGTCTCATCAGAGGTGGCCACCGATGGCCCAAGCAGCAGCAGCAGCAGCAGCAGCAG 145
QY 223 gccagagcgtggttcccaaacccacgagcggggtggtggtggtggtggtggtggtggtggt 282
DB 146 GCCAGAGCGGTGGTTCCTCCAAACCCAGCAGCGGGCGCGAGTATGTGTGGCCAGCTGGAT 205
QY 283 gacctgtcaactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 342
DB 206 GACCTCGTCAACTGGGCGCGCGGAGTCTCTGTGGCCCATGACCTTCGCGCTGGCGCTGC 265
QY 343 tgcgctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 402
DB 266 TGCCTGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
QY 403 gttctccggtcagcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggc 462
DB 326 GTCTTCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 385
QY 463 aagatggtcccgagcgttgcgaaggttgcgaaggttgcgaaggttgcgaaggttgcgaaggttgc 522
DB 386 AAGATGGCCCGCGCGCTTCGCAAGGTCTACGACAGATGCGGAGCGCGCTACGCTGGTC 445
QY 523 tccatgggagctgcgcacgagagaggttaccatttcttacttcttacttcttacttcttacttct 582
DB 446 TCCATGGGAGCTGGCGCAACGGAGGAGGTACTTACCACATTCTCTACTCTCTCTCTCTCTCT 505
QY 583 ggtctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 628
DB 506 GGCTCGACCGCATGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551

RESULT 14
BG419411 910 bp mRNA EST 14-MAR-2001
LOCUS BG419411
DEFINITION 602446102F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4584528 5',

mRNA sequence.
 BG419411
 VERSION BG419411.1 GI:13325917
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 910)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI310 row: d column: 01
 High quality sequence stop: 702.
 Location/Qualifiers
 1..910
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4584528"
 /clone_lib="NIH_MGC_14"
 /tissue_type="renal cell adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
 ECORI; CDNA made by oligo-dt priming. Directionally
 cloned into EcorI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size: selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 195 a 279 c 278 g 156 t
 ORIGIN
 Query Match 61.4%; Score 506; DB 11; Length 910;
 Best Local Similarity 99.8%; Pred. No. 3.8e-133;
 Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 72 gccaagatggcggtgtcagctcctgctgcgcgcgtcccgatccttgggtctgcg 131
 DB 1 GGCACAGATGGCGGTGTGTGTCAGCTCTGCTGCGGCTCCGATCCTTGGTCTGCG 60
 QY 132 ctccagcgtggccggcgtgtcagcagcagaggtgtccatcagagcgtggccaccgatgg 191
 DB 61 CTCAGAGGTGGCGTGTGTCAGGACGAGGTGTCCATCAGAGGTGGCCACCGATGG 120
 QY 192 cccaagcagcaccagcctccctgcacagcagcagcgcgtggctcccaaccacagcag 251
 DB 121 CCCAAGCAGCACCCGCTGCTGCCCTGCCAAGGCCAGAGCGTGGCTCCCAACCCAGCAG 180
 QY 252 ccggggggagatgtggtggccaaagtggatgacctcgtcaactggccgcgcgcggagttc 311
 DB 181 CCGGGCGAGTATGTGGTGCCAACTGGATGACCTCGTCAACTGGCGCGCGGAGTTC 240
 QY 312 tctgtgcccataccttggcctgctcgtcgcgtggagatgatcacatggcagc 371
 DB 241 TCTGTGGCCCATGACCTTCGGCTGGCTGCTGCGCGTGGAGATGATGACATGGCAGC 300
 QY 372 acccgcgtacacatgacacgcttttggcgtgggtcttccgcgcgcgcgcgcgcagtcoga 431
 DB 301 ACCCGCTACGACATGACCGCTTTGGCGTGTGCTTCCGCGCGCAGCCGCGCATCCGA 360
 QY 432 cgtcatatcgtggcggcacactcaccacacacagatggccccagcgttcgcaaggtcta 491

Db 361 CGTCATGATCGTGGCGGCACACTCACCACAAAGATGGCCCGAGCGCTTCGCAAGGTCTA 420
 QY 492 cgaccagatcccgagcgcgcgtacgtgtctctccatgggagcgtgcgcaacggagggag 551
 Db 421 CGACCAGATCGCGAGCGCGGTACGTTCTCCATGGGGAGTGGCGCCACCGGAGGAGG 480
 QY 552 ctactaccactattcctactcgtgtgtgagggcgtgcgaccacatcgtcccgatgacat 611
 Db 481 CTACTACCATTTCTACTCGTGTGTGAGGGGCTCGACCGCATCGTCCGCGCGTGACAT 540
 QY 612 ctatcatcccgaggctgcg 628
 Db 541 CTACATCCCGAGGTGCC 557
 RESULT 15
 BG491103
 LOCUS 738 bp mRNA EST 27-MAR-2001
 DEFINITION G02519937F1 NTH_MGC_18 Homo sapiens cDNA clone IMAGE:4638441 5',
 mRNA sequence.
 ACCESSION BG491103
 VERSION BG491103.1 GI:13452615
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 738)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI400 row: j column: 10
 High quality sequence stop: 738.
 Location/Qualifiers
 1..738
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4638441"
 /clone_lib="NIH_MGC_18"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
 ECORI; CDNA made by oligo-dt priming. Directionally cloned
 into EcorI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."
 BASE COUNT 125 a 252 c 235 g 126 t
 ORIGIN
 Query Match 61.3%; Score 505; DB 11; Length 738;
 Best Local Similarity 99.8%; Pred. No. 8.4e-133;
 Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 73 gccaagatggcggtgtgtcagctcctgctgcgcgcgtcccgatccttgggtctgcgc 132
 Db 1 GGCACAGATGGCGTGTGTGTCAGCTCTGCGGCTCCGATCCTTGGTCTGCGC 60
 QY 133 tcacagcgtggccgcgcgtgtgagcagcagcagcagcagcagcagcagcagcagcagcagc 192
 Db 61 TCCAGCGTGGCGCTGGCTGTGTCAGGCGCAGGAGTGTCCATCAGAGCGTGGCCCGCATGCG 120

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 20:29:55 ; Search time 33.73 Seconds
(without alignments)
481.031 Million cell updates/sec

Title: US-09-525-867-1
Perfect score: 1118
Sequence: 1 MAVLSAPGLRGRILGRSS.....ILQLQKIKRRRIQIWR 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965.5	86.4	216	2 S22371	NADH dehydrogenase
2	736	65.8	215	2 S48826	NADH dehydrogenase
3	732.5	65.5	213	2 T07603	NADH dehydrogenase
4	731	65.4	213	2 T44818	NADH dehydrogenase
5	726.5	65.0	199	2 T26329	hypothetical prote
6	725	64.8	218	2 S22286	NADH dehydrogenase
7	718	64.2	182	2 S78166	NADH dehydrogenase
8	654.5	58.5	174	2 F71692	NADH dehydrogenase
9	654	58.5	191	2 D82821	NADH-ubiquinone ox
10	626	56.0	173	2 E42573	psbG homolog prote
11	613	54.8	160	2 H81221	NADH dehydrogenase
12	605	54.1	160	2 F81922	NADH dehydrogenase
13	561	50.2	156	1 F3PPG	psbG protein - Par
14	521	46.6	179	2 C70413	NADH dehydrogenase
15	497.5	44.5	159	2 E64677	NADH dehydrogenase
16	496.5	44.4	159	2 C71838	NADH dehydrogenase
17	474	42.4	181	2 C75388	NADH dehydrogenase
18	461.5	41.3	167	2 H81252	NADH dehydrogenase
19	457	40.9	181	2 T11899	NADH dehydrogenase
20	444	39.7	220	1 E65000	NADH dehydrogenase
21	444	39.7	220	2 D85859	NADH dehydrogenase
22	443	39.6	184	2 T34623	NADH dehydrogenase
23	441.5	39.5	235	2 F84221	NADH dehydrogenase
24	438	39.2	248	1 S04437	NADH dehydrogenase
25	437.5	39.1	184	2 C70647	probable nuob prot
26	432	38.6	225	2 C83316	NADH dehydrogenase
27	431.5	38.6	247	1 F2NTG	psbG protein - com
28	428	38.3	224	2 A84948	NADH dehydrogenase
29	427.5	38.2	246	1 F2R2G	psbG protein - ric

30	426.5	38.1	243	1 F2LVG	psbG protein - liv
31	424.5	38.0	219	1 F2YB32	hypothetical ndhk
32	423.5	37.9	226	2 T09637	NADH dehydrogenase
33	423.5	37.9	245	2 S09666	psbG protein - whe
34	422.5	37.8	248	1 F2ZMG	NADH dehydrogenase
35	420.5	37.6	245	2 S74216	NADH dehydrogenase
36	408.5	36.5	224	2 S27978	NADH dehydrogenase
37	350.5	31.4	202	2 C72621	probable NADH dehy
38	344	30.8	96	2 PX0053	NADH dehydrogenase
39	319.5	28.6	353	2 D69478	NADH dehydrogenase
40	311	27.8	195	2 A75114	NADH dehydrogenase
41	301	26.9	195	2 D71019	probable NADH-ubiq
42	295	26.4	178	2 D72281	hypothetical prote
43	283	25.3	173	2 E71017	probable NADH-ubiq
44	281	25.1	170	2 D75115	co-induced hydroge
45	275	24.6	255	2 S08625	hydrogenase (EC 1.

ALIGNMENTS

RESULT 1

S22371
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 26-Aug-1999
C:Accession: S22371
R:Arizmendi, J.M.; Runswick, M.J.; Skehel, J.M.; Walker, J.E.
FEBS Lett. 301, 237-242, 1992

A:Title: NADH: ubiquinone oxidoreductase from bovine heart mitochondria. A fourth m
A:Reference number: S22371; MUID:92249573
A:Accession: S22371
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-216 <AR>
A:Cross-references: EMBL:X65020; NID:gl1256; PIDN:CAA46154.1; PID:9599691
C:Superfamily: psbG protein
C:Keywords: NAD; oxidoreductase

Query Match 86.4%; Score 965.5; DB 2; Length 216;
Best Local Similarity 85.7%; Pred. No. 2.6e-81;
Matches 186; Conservative 11; Mismatches 15; Indels 5; Gaps 2;

QY	1	MAVLSAPGLRGRILGRSSVGPVAVQARGVHQSVATDGPSTQPALPKARAVAPK----	P 56
Db	1	MAALAALRL-HPILAVRSGVGAALQVRGVHSSMAADSPSTQPAVSOARAVVEKPAALP	59
QY	57	SRGEYVYVAKLDDLNVNARRSSLWPMTGLACCAVEMMHMAAPRYDMDRFGVVFASPRQ	116
Db	60	SRGEYVYVAKLDDLNVNARRSSLWPMTGLACCAVEMMHMAAPRYDMDRFGVVFASPRQ	119
QY	117	SDVMIVAGTLTKMAPALKVYDQMPPEPVYVSMGSCANGGGYHYYSVYRGCDRIYVP	176
Db	120	SDVMIVAGTLTKMAPALKVYDQMPPEPVYVSMGSCANGGGYHYYSVYRGCDRIYVP	179
QY	177	DIYFGCPPTAEALLYGILQKIKRRRIQIWR 213	
Db	180	DIYFGCPPTAEALLYGILQKIKRRRIQIWR 216	

RESULT 2

S48826
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST precursor - wild cabbage
C:Species: Brassica oleracea (wild cabbage)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Mar-2001
C:Accession: S48826
R: Pogson, B.J.; Downs, C.G.; Davies, K.M.; Morris, S.C.; Buchanan-Wollaston, V.
submitted to the EMBL Data Library, September 1994
A:Description: Nucleotide sequence of a cDNA clone from broccoli (Brassica oleracea
A:Reference number: S48827
A:Accession: S48826
A:Molecule type: mRNA

QY Z/ ARGVHQSVATDGPSSSTQPALPKARAVAP- - - KPSRRGEIVVAKDDLVNWARSSLEWFM 8Z

psbc homolog protein - Paracoccus denitrificans
C:Species: Paracoccus denitrificans
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Nov-1998
C:Accession: E42573
R:Xu, X.; Matsuno-Yagi, A.; Yagi, T.
Biochimie 31, 6925-6932, 1992
A:Title: Gene cluster of the energy-transducing NADH-quinone oxidoreductase of Paracoccus
A:Reference number: A42573; MUID:92345253
A:Accession: E42573
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <XU1>
A:Experimental source: ATCC 13543
A:Note: sequence extracted from NCBI backbone (NCBIN:110121, NCBIP:110125)

F01992
NADH dehydrogenase (ubiquinone) [EC 1.6.5.3] I chain B NMA0018 [imported] - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: F01992
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Holtzoyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandram, R.; Whitehead, P.M.A.
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 224c/p1

A:Reference number: A81775; MUID:20222556
A:Accession: F81992
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <PAR>
A:Cross-references: GB:AL157959; NID:g7378778; PIDN:CAB83338.1; PID:g737879
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: nuoB; NMA0018
C:Superfamily: psbG protein
C:Keywords: NAD; oxidoreductase

Query Match 54.1%; Score 605; DB 2; Length 160;
Best Local Similarity 73.2%; Pred. No. 2.3e-48;
Matches 104; Conservative 22; Mismatches 16; Indels 0; Gaps 0;

QY 62 YVYAKLDDLNVNARRSLPMTFGLACCAVEMHMAAPRYDMDRFGVVFRASPROSDVMI 121
DB 11 FITTSADTVLNYMRTGSLWPVTFGLACCAVEMHAGWARYDLDLDFGIIFRPSRQADLMI 70
QY 122 VAGTLNKMALPARKYVDQMPERPVVYVSMGSCANGGGYHYHSYSVVRGCDRIYVPDVIYP 181
DB 71 VAGTLNKMASALRRYVDQAEPRVLSMGSCANGGGYHYHSYSVVRGADRVVPDVIYP 130
QY 182 GCPPTAEALLYGILQLOQKIKR 203
DB 131 GCPPTAEALLYGILQLOQKIKR 152

RESULT 13
F2PPG
psbG protein - Paramesium tetraurelia mitochondrion
C:Species: mitochondrion Paramesium tetraurelia
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 19-Jan-2001
C:Accession: S07736; JS0235
R:Pritchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummings, D.J.
Nucleic Acids Res. 18, 173-180, 1990
A:Title: Nucleotide sequence of the mitochondrial genome of Paramesium.
A:Reference number: S07725; MUID:90174913
A:Accession: S07736
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-156 <PRI>
A:Cross-references: EMBL:X15917; NID:g13256; PIDN:CAA34045.1; PID:g578753
R:Pritchard, A.E.; Venuti, S.E.; Ghalambor, M.A.; Sable, C.L.; Cummings, D.J.
Gene 78, 121-134, 1989
A:Title: An unusual region of Paramesium mitochondrial DNA containing chloroplast-like
A:Reference number: JS0231; MUID:89357489
A:Accession: JS0235
A:Molecule type: DNA
A:Residues: 1-122, '126, 'G, 128-156 <PRI2>
A:Cross-references: GB:M26930; NID:g341550; PIDN:AAA79257.1; PID:gi019632
A:Experimental source: strain sp. 4.51
A:Note: the source is designated as Paramesium aurelia species 4 stock 51, now designated
C:Genetics:
A:Gene: psbG
A:Genome: mitochondrion
A:Genetic code: SGC6
A:Start codon: ATA
C:Superfamily: psbG protein
C:Keywords: mitochondrion

Query Match 50.2%; Score 561; DB 1; Length 156;
Best Local Similarity 63.4%; Pred. No. 2.4e-44;
Matches 92; Conservative 33; Mismatches 20; Indels 0; Gaps 0;

QY 59 RGEYVYAKLDDLNVNARRSLPMTFGLACCAVEMHMAAPRYDMDRFGVVFRASPROSD 118
DB 4 KADFLKLSANLISWARGSGFWLPTFLGACCAEMHATVSYRDFDRFGVIFRATPQAD 63
QY 119 VMIVAGPLTNKMAPALRKVYDQMPERPVVYVSMGSCANGGGYHYHSYSVVRGCDRIYVPDI 178

```
QY 63 VVAKLDDLVNWARSSSLWPMTEGLACCAVEMMHMAAPRYDMDREGVVERASPROSDVNIY 122
Db 7 VLSLTKLLNWGSNSLWPEYGLACCAIEMMATGGSRFDFRGTIFRASPROSDVNIY 66
QY 123 AGTLTKMAPALKVYDQMPPEPRYVVMGSCANGGYYHYSYVVRGCDRIVPVDIYIPG 182
Db 67 AGTLTKKHAEFMRRLYDQMPPEPKWVISMGCANTGGMEN-TYATVOGADRVVPVDIYLPG 125
QY 183 CPPTAEALLYGILQLOKTKRERRLQ 208
Db 126 CAPRPETLQYALWLODKIRSKAIK 151
```

Search completed: february 12, 2002, 20:45:24
Job time: 929 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 16:55:05 ; Search time 117.26 Seconds
(without alignments)
6024.527 Million cell updates/sec

Title: US-09-525-867-9
Perfect score: 824
Sequence: 1 cggctcagcgctcagcg.....cgggcaaaaaaaaaaaaaa 824

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	750.8	91.1	772	22	AAI58780 Human polynucleoti
2	738.8	89.7	752	22	AAI60566 Human polynucleoti
3	715.4	86.8	785	19	AAAC59553 Human secreted pro
4	384.6	46.7	789	19	AAV61275 3' CDNA sequence o
5	384.6	46.7	789	19	AAV58509 3' fragment of pro
6	384.6	46.7	789	21	AAA06272 Human immunogenic
7	384.6	46.7	789	22	AAI10031 Human prostate tum
8	384.6	46.7	789	22	AAH93388 Human prostate-spe
9	384.6	46.7	789	22	AAH84702 Human prostate-spe
10	384.6	46.7	789	22	AAH02453 Prostate tumour an
11	279.2	33.9	911	21	AAAC43639 Zea mays DNA fragm

12	251.8	30.6	937	21	AAAC36686 Arabidopsis thalia
13	251.8	30.6	976	21	AAAC38037 Arabidopsis thalia
14	237.8	28.9	1266	21	AAAC41785 Arabidopsis thalia
15	208.6	25.3	72750	21	AAAB1468 N. meningitidis pa
16	208.6	25.3	349980	21	AAAF21544 Neisseria meningit
17	208.6	25.3	1437668	21	AAAB1490 N. meningitidis B
18	145.2	17.6	458	21	AAAC40995 Zea mays DNA fragm
19	124.4	15.1	446	21	AAAC95299 Cat flea head and
20	112.2	13.6	58909	22	AAAF28543 Genomic fragment #
21	92.4	9.9	134525	11	AAAT4101 Human gene signatu
22	81.8	9.9	134525	11	AAAF04525 Total base sequenc
23	52.4	6.4	2322	21	AAAB1961 Modified HIV-1 Env
24	52.4	6.4	2541	21	AAAB1950 Modified HIV-1 Env
25	52.4	6.4	2541	21	AAAB1952 Modified HIV-1 Env
26	50.8	6.2	2322	21	AAAB1959 Modified HIV-1 Env
27	50.8	6.2	2322	21	AAAB1967 Modified HIV-1 Env
28	50.8	6.2	2541	21	AAAB1951 Modified HIV-1 Env
29	49.4	6.0	1950	22	AAAF5384 Codon-optimised mu
30	49.2	6.0	2322	21	AAAB1960 Modified HIV-1 Env
31	49.2	6.0	2541	21	AAAB1953 Modified HIV-1 Env
32	47.6	5.8	2088	12	AAQ14479 Truncated HSVgB ge
33	47.6	5.8	2925	12	AAQ14455 HSV surface antige
34	47.6	5.8	3465	12	AAQ14478 HSVgB gene. Herpe
35	46.8	5.7	672	22	AAAF9964 Partial nucleotide
36	46.2	5.6	3642	6	AAAF50364 DNA fragment encod
37	46	5.6	1323	21	AAAF70432 HIV gp120.modSF162
38	46	5.6	1387	21	AAAF70431 HIV gp120.modSF162
39	46	5.6	1453	21	AAAF70430 HIV gp120.modSF162
40	46	5.6	1836	21	AAAF70438 HIV gp140.mut7.mod
41	46	5.6	1836	21	AAAF70441 HIV gp140.mut8.mod
42	46	5.6	1836	21	AAAF70444 HIV gp140.modSF162
43	46	5.6	1944	21	AAAF70434 HIV gp140.modSF162
44	46	5.6	1944	21	AAAF70435 HIV gp140.mut7.mod
45	46	5.6	1944	21	AAAF70437 HIV gp140.mut8.mod

ALIGNMENTS

RESULT 1
AAI58780
ID AAI58780 standard; cDNA; 772 BP.
XX
AC AAI58780;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 983.
XX
DE Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; SS.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0695036.
PR 29-NOV-2000; 2000US-0727344.
XX

CC Alzheimer's, Parkinson's disease, Huntington's disease, ankyrotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ sequence 752 BP; 130 A; 260 C; 236 G; 136 T; 0 other;

Query Match	89.7%	Score 738.8	DB 22	Length 752
Best Local Similarity	99.1%	Pred. No. 7.4e-153		
Matches 743	Conservative 0	Mismatches 7	Indels 0	Gaps 0
QY	75	caagatgcgcggtgctgcagctctcgccctgcgcggcttcgcgatacttggtctgcgctc	134	
Db	1	caagatgcgcggtgctgcagctctcgccctgcgcggcttcgcgatacttggtctgcgctc	60	
QY	135	cagcgtggcgccggctgtgcaggcacagagtgtccatcacagcgtggccacacgatggccc	194	
Db	61	cagcgtggcgccggctgtgcaggcacagagtgtccatcacagcgtggccacacgatggccc	120	
QY	195	aagcagcaccacagcctgccttgccaaaggccagacgcgtggctccaaacccacagacgcg	254	
Db	121	aagcagcaccacagcctgccttgccaaaggccagacgcgtggctccaaacccacagacgcg	180	
QY	255	ggcgagagtatggtggccaaagctggatgacctgctcaactggcgccgcgcgaagtctct	314	
Db	181	ggcgagagtatggtggccaaagctggatgacctgctcaactggcgccgcgcgaagtctct	240	
QY	315	gtggcccatgaccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc	374	
Db	241	gtggcccatgaccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc	300	
QY	375	cgcgtacacatggaacgccttggcgtgggttcttcgcgcgcgcgcgcgcgcgcgcgcgcgc	434	
Db	301	cgcgtacacatggaacgccttggcgtgggttcttcgcgcgcgcgcgcgcgcgcgcgcgcgc	360	
QY	435	catgatcgtggccgcgcacactcaacacaaagatggcccccagcgttcgcgaaggtctacga	494	
Db	361	catgatcgtggccgcgcacactcaacacaaagatggcccccagcgttcgcgaaggtctacga	420	
QY	495	ccagatgcccggagccgcgcctacgttggtctccatggggagctgcgcgaagggaggctca	554	
Db	421	ccagatgcccggagccgcgcctacgttggtctccatggggagctgcgcgaagggaggctca	480	
QY	555	ctaccacattctactcgttggtggagggtgcgcacgcacatcgtcccgtagacatcta	614	
Db	481	ctaccacattctactcgttggtggagggtgcgcacgcacatcgtcccgtagacatcta	540	
QY	615	catcccgagctgccacactacgcgcgagccctgctctacggcactcctgcagctgcagag	674	
Db	541	catcccgagctgccacactacgcgcgagccctgctctacggcactcctgcagctgcagag	600	
QY	675	gaagatcaagcgggagcggagctgcagatctgttacccgcaggtagcgcgcgcgcgcgcgcg	734	
Db	601	gaagatcaagcgggagcggagctgcagatctgtgtacccgcaggtagcgcgcgcgcgcgcgcg	660	
QY	735	cgcgcgcgcctctgcgcgcctctgtccccacgcctgtgtgtcccgtagggttgtcaata	794	
Db	661	cgcgcgcgcctctgcgcgcctctgtccccacgcctgtgtgtcccgtagggttgtcaata	720	
QY	795	aacctgcccctggggcaaaaaa	824	
Db	721	aacctgcccctgggctggcgcacaaaaa	750	

RESULT	3
AAC59553	
ID	AAC59

AAC59553;
15-FEB-2001 (first entry)
Human secreted protein cDNA sequence #47.
Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
neurological disease; infection; human; secreted protein; ss.
Homo sapiens.
WO2000055352-A2.
21-SEP-2000.
09-MAR-2000; 2000WO-US06044.
12-MAR-1999; 99US-0124099.
03-DEC-1999; 99US-0168661.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Ruben SM, Komatsoulis G;
WPI: 2000-602124/57.
P-PSDB; AAB34261.
Novel human secreted proteins useful for diagnosis, prevention and
treatment of disorders including neurological, cell proliferative,
cardiovascular, autoimmune and inflammatory disorders and microbial
infections -
Claim 1; Page 345; 383pp; English.
The invention relates to the isolation of genes AAC59507-C59556 encoding
50 human secreted proteins AAB34218-B34264). The genes can be used to
generate fusion proteins by linking to the gene for the human
immunoglobulin G Fc portion (AAC59498) for increasing the stability of
the fusion protein as compared to the human protein only. The genes and
proteins are useful for preventing, ameliorating or treating medical
conditions, e.g. by protein or gene therapy. The genes are isolated
from a range of human tissues disclosed in the specification. The
nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
and ovarian cancer, and other cancers of the adrenal gland, bone, bone
marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
(b) immune disorders e.g. Addison's disease, allergies, autoimmune
haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
disease, multiple sclerosis, rheumatoid arthritis and ulcerative
colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
wound healing; (e) neurological diseases e.g. cerebral anoxia and
epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
and parasitic infections.
Sequence 785 BP; 173 A; 255 C; 236 G; 121 T; 0 other;

	Query Match	86.8%	Score 715.4	DB 21	Length 705
	Best Local Similarity	99.7%	Pred. NO. 9.8e-148		
	Matches 727	Conservative 0	Mismatches 1	Indels 1	Gaps 1
Qy	96	tcttgccctgcgcgctctccggatccctgctcgcgtccagcgtggccgcgctgtgca	155		
Db	23	tcttgccctgcgcgctctccggatccctgctcgcgtccagcgtggccgcgctgtgca	82		
Qy	156	ggcacgagtgctccatcagacgctggccacgctggccacgacgacacccagcctgacct	215		
Db	83	ggcacgagtgctccatcagacgctggccacgctggccacgacgacacccagcctgacct	142		

RESULT	3
AAC59553	
ID	AAC59


```
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162144.

Query Match      58.3%; Score 652; DB 21; Length 143;
Best Local Similarity 80.4%; Pred. No. 4.6e-65;
Matches 115; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 71 VNWARSSLPMTFTGLACCAVEMHMAAPRDMDRFGVVFRRASPRQSDVMTVAGTTLNKM 130
Db 1 mwwartgs.wpmftglaccavemhmtgaaryldrfgrfprsqsdcm.vagtlnkm 60

QY 131 APALRVYDQMPPEPRVVSVMGSCANGGYHYHSYVWVGCDRIYFVDIYIFGCPPTAEAL 190
Db 61 apalrvyqdmpeprvwismgscangcgyhyhsyvsvrgcdriypvdiypvgcptaeal 120

QY 191 LYGILQKRIKRRKQLQWYRR 213
Db 121 lygllqlgklnrkrdflhwk 143

RESULT 15
AAG28058
ID AAG28058 standard; Protein; 132 AA.
XX AC AAG28058;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33133.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123168.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
```

```
PR 13-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 03-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 602; DB 21; Length 132;
Best Local Similarity 81.8%; Pred. No. 1.7e-59;
Matches 108; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 82 MTFGLACCAVEMHMAAPRYDMDREGVVFRRASPROSDVMIVAGTLTNKMAPALRKVYDQM 141
Db 1 mtfglaccavemhmtgaarydldrfgilfrpsrgscdmivagtltnkmapalrkvydqm 60

QY 142 PPRYVVMGSCANGGGYHYHSYVVRGCDRIVPVDIYIPGCPPTAEALLYGILQLQKRI 201
Db 61 pprwvismgscangcgvyhysvrgcdrivpvdiypgcpptaeallygllqgkxi 120

QY 202 KRERLQIWYR 213
Db 121 nrkdfhlwvwnk 132

Search completed: February 12, 2002, 20:37:07
Job time: 2457 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 17:47:55 ; Search time 102.3 Seconds
(without alignments)
6905.533 Million cell updates/sec

Title: US-09-525-867-9
Perfect score: 824
Sequence: 1 cggtcgagcggtcgagcg.....cgggcaaaaaaaaaaaaaa 824

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_1101.*

1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	747	90.7	772	AA158780	Human polynucleoti
2	735	89.2	752	AA160566	Human polynucleoti
3	592	71.8	785	AA160566	Human secreted pro
4	40	4.9	117	AA124101	Human gene signatu
5	27	3.3	789	AA161275	3' CDNA sequence o
6	27	3.3	789	AA158509	3' fragment of pro
7	27	3.3	789	AA160272	Human immunogenic
8	27	3.3	789	AA10031	Human prostate tum
9	27	3.3	789	AA153388	Human prostate-spe
10	27	3.3	789	AA184702	Human prostate-spe
11	27	3.3	789	AA102453	Prostate tumour an

12	27	3.3	1038	21	AA249556	Partial coding reg
13	27	3.3	1989	20	AA26147	Rat pan-s/tyk recep
14	25	3.0	381	21	AA132035	Plant microsatelli
15	24	2.9	1780	21	AA121665	Human breast and o
16	24	2.9	1804	22	AA157575	Human brain cell s
17	23	2.8	1107	21	AA189464	LeuB subunit of 3-
18	23	2.8	2037	22	AA151958	Mycobacterium tube
19	22	2.7	436	21	AA141443	Zea mays DNA fragm
20	22	2.7	597	22	AA187764	Peppermint plant o
21	22	2.7	631	22	AA107813	Human cDNA clone (
22	22	2.7	1360	22	AA142475	cDNA sequence of t
23	22	2.7	1859	22	AA125214	Maize ferulate-5-h
24	22	2.7	1859	22	AA105736	Maize ferulate-5-h
25	22	2.7	1918	22	AA114977	Human cDNA sequenc
26	22	2.7	2174	22	AA186635	Human cDNA sequenc
27	22	2.7	2308	15	AA106638	Human Activin rece
28	22	2.7	2308	20	AA109845	Human hALK-5 clone
29	22	2.7	2308	21	AA148863	Human Transforming
30	22	2.7	2600	17	AA143624	Chromatin regulato
31	22	2.7	5216	22	AA128355	Nucleotide sequenc
32	21	2.5	152	21	AA104711	Human secreted pro
33	21	2.5	273	21	AA141424	Zea mays DNA fragm
34	21	2.5	374	21	AA131318	Plant microsatelli
35	21	2.5	379	21	AA131636	Plant microsatelli
36	21	2.5	385	21	AA151712	Zea mays DNA fragm
37	21	2.5	393	22	AA181463	Rice glutaredoxin
38	21	2.5	420	21	AA141346	Zea mays DNA fragm
39	21	2.5	424	21	AA151744	Zea mays DNA fragm
40	21	2.5	428	19	AA138430	cDNA encoding the
41	21	2.5	456	21	AA139822	Zea mays DNA fragm
42	21	2.5	459	19	AA138429	cDNA encoding the
43	21	2.5	472	21	AA140956	Zea mays DNA fragm
44	21	2.5	473	21	AA143526	Zea mays DNA fragm
45	21	2.5	482	21	AA143561	Zea mays DNA fragm

ALIGNMENTS

RESULT	1
AA158780	
ID	AA158780 standard; cDNA; 772 BP.
AC	AA158780;
XX	
XX	
DT	22-OCT-2001 (first entry)
XX	
XX	Human polynucleotide SEQ ID NO 983.
DE	
KW	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0651450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 752 BP; 130 A; 260 C; 236 G; 126 T; 0 other;

Query Match 89.2%; Score 735; DB 22; Length 752;
Best Local Similarity 100.0%; Pred. No. 3.5e-285;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 caagatggcgggtgctgcagctccctggcgtgcggcgttcctcggtacccctgtgtcgtcgctc 134
|||||
Db 1 caagatggcgggtgctgcagctccctggcgtgcggcgttcctcggtacccctgtgtcgtcgctc 60
QY 135 cagctggcggcgggtgctgcagcagcagaggtgtccatcagagcgtggccacccgatggccc 194
|||||
Db 61 cagctggcggcgggtgctgcagcagcagaggtgtccatcagagcgtggccacccgatggccc 120
QY 195 aagcagcaccacgcctgcctcccaaggccagagccgtggctcccaaccacccagcgcg 254
|||||
Db 121 aagcagcaccacgcctgcctcccaaggccagagccgtggctcccaaccacccagcgcg 180
QY 255 gggcagatgtgtgtggccaaagtgtgacacctgtcaactggcccccggagttctct 314
|||||
Db 181 gggcagatgtgtgtggccaaagtgtgacacctgtcaactggcccccggagttctct 240
QY 315 gtggcccatgaacctggcgtgcctgtgcctgtgcctgtgagatgatgcacatggcagcacc 374
|||||
Db 241 gtggcccatgaacctggcgtgcctgtgcctgtgcctgtgagatgatgcacatggcagcacc 300
QY 375 ccgtacagatggacccgttttggcgtgtgttcctgcgcgcagccgcgcagtcgcagct 434
|||||
Db 301 ccgtacagatggacccgttttggcgtgtgttcctgcgcgcagccgcgcagtcgcagct 360
QY 435 catgctgtggcggccgacactcaccacaagatggccacagcgttcgcaaggtctacga 494
|||||
Db 361 catgctgtggcggccgacactcaccacaagatggccacagcgttcgcaaggtctacga 420
QY 495 ccagatgcggagcggcgtactgtgttccatggggagcgtgcgccaacggagaggcta 554
|||||
Db 421 ccagatgcggagcggcgtactgtgttccatggggagcgtgcgccaacggagaggcta 480
QY 555 ctaccactattcctcgtgtgtgaggggtgcgaccgacatcgtgccgtggacatcta 614
|||||
Db 481 ctaccactattcctcgtgtgtgaggggtgcgaccgacatcgtgccgtggacatcta 540
QY 615 catccagagctgccacactacggcggagcgcctgtctctacggcatcctgcagctgcagag 674
|||||
Db 541 catccagagctgccacactacggcggagcgcctgtctctacggcatcctgcagctgcagag 600
QY 675 gaagatcaagcgggagcggagcgtcagatgtgttacgcagggtagcgcgcgccccgcgg 734
|||||
Db 601 gaagatcaagcgggagcggagcgtcagatgtgttacgcagggtagcgcgcgccccgcgg 660
QY 735 ccgcggagcctgtgcctcctctgtccacagcctgtgttccctgtgaggttgtcaata 794
|||||
Db 661 ccgcggagcctgtgcctcctctgtccacagcctgtgttccctgtgaggttgtcaata 720
QY 795 aacctgcctcctgggc 809
|||||
Db 721 aacctgcctcctgggc 715

RESULT 3
AAC59553
ID AAC59553 standard; cDNA; 785 BP.

XX AAC59553;
AC 15-FEB-2001 (first entry)
DT Human secreted protein cDNA sequence #47.
DE
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ss.
XX Homo sapiens.
OS
XX WO200055352-A2.
PN
XX 21-SEP-2000.
PD
XX 09-MAR-2000; 2000WO-US06044.
PF
XX 12-MAR-1999; 950S-0124099.
PR
XX 03-DEC-1999; 950S-0168661.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-502124/57.
DR P-PSDB; AAB34261.
XX
PT Novel human secreted proteins useful for diagnosis, prevention and
PT treatment of disorders including neurological, cell proliferative,
PT cardiovascular, autoimmune and inflammatory disorders and microbial
PT infections
XX
PS Claim 1; Page 345; 383pp; English.
XX
CC The invention relates to the isolation of genes AAC59507-C59556 encoding
CC 50 human secreted proteins AAB34218-A34264. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (AAC59498) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX Sequence 785 BP; 173 A; 255 C; 236 G; 121 T; 0 other;

Query Match 71.8%; Score 592; DB 21; Length 785;
Best Local Similarity 99.9%; Pred. No. 5.9e-228;
Matches 712; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 96 tcttgccctgcggcgttcctcggtacccctgtgtcgtccagcgtggcccggtgtgca 155
|||||
Db 23 tcttgccctgcggcgttcctcggtacccctgtgtcgtccagcgtggcccggtgtgca 82
QY 156 ggcacgaggtgtccatcagcgtggccacccagcgtggccacccacccacccgcct 215
|||||
Db 83 ggcacgaggtgtccatcagcgtggccacccagcgtggccacccacccacccgcct 142

XX WPI; 2001-425873/45.
DR
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
PS Claim 1; Page 240; 543pp; English.
CC
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC i cells prepared using (I) or (II) are used treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

Query Match 3.3%; Score 27; DB 22; Length 789;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 722 gcgcgcccgcgcgcggagcctgt 748
|||||
Db 151 GCCGCCGCGCGCGCGAGCCTGT 125

RESULT 10
AAH84702/c
ID AAH84702 standard; CDNA; 789 BP.
AC AAH84702;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human prostate-specific 3' cDNA sequence J1-21.
XX
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW prostate specific antigen; PSA; ss.
XX
OS Homo sapiens.
XX
PN WO200134802-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30904.
XX
PR 12-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAM, Wang A;
XX
WI; 2001-308785/32.
XX
PT Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate-specific protein, useful in the diagnosis and therapy of
PT prostate cancer -
XX
PS Claim 31; Page 140; 325pp; English.

CC New markers from the chromosomes of the human genome can be isolated and
CC positional information of isolated sequences can be provided, for use in

Search completed: February 12, 2002, 19:56:04
Job time: 7689 sec

Qy 141 gggccggtgtgcagcacgagggtgtccatcagagcggtgccaccgatggcccaagcag 200
||| ||| | ||| | ||| | ||| | ||| |
Ddb 739 GGGGATGGTGNNAAGSANGNGTTTCNTAGAGNNGGCCACCGGGGGCCCAAGGG 680


```
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      10.98; Score 89.6; DB 4; Length 4403765;
Best Local Similarity 59.8%; Pred. No. 2.4e-11;
Matches 292; Conservative 0; Mismatches 190; Indels 6; Gaps 2;

QY 207 gctgcctcgcacaaaggccagagccgtggtcccaaccacccagcagccggcgagtgatgt 266
Db 3507485 gccgacgtgggattgaggtaggcgctgggactggagaaacagctgcccgccggatcct 3507544

QY 267 gctgcacagctgagatgacctcgtcaactggccggcgagttctctgtggcccatgac 326
Db 3507545 gctgtcacctgcagaaggctggcggtctatgtccgcaaaaactcctgtggccggcaac 3507604

QY 327 ctccggcctgcctgctgcgcgtggagatgacacatgacagcaccgcccgctacacat 386
Db 3507605 attcgattggcgtgctgtgcgcatcagatgctggcagccgggacaaagtttgacat 3507664

QY 387 ggaacgcgtttggcggtg---ttctccgcgcagccgcgcagctccgagctcagatcgt 443
Db 3507665 tgcgcgttccggtggaacggttctcgcgcagccgcgcagcgagcagatctgatcgt 3507724

QY 444 ggcgcgcacacacccaaagatggcccgagcttcgcagcgttcgcaagttcagacagatgcc 503
Db 3507725 ggcgcgcgcgcgcagcagagatggcgcgttactgcgcgcagatctatgacagatggc 3507784

QY 504 ggaacgcgcgtcagctgggtctccatgggagcgtgcgcacacggagagctactaccacaa 563
Db 3507785 ggaacgcgaatgggttctggcgaatgggtgtgtgcgcctcgtcagtgaggattcaacaa 3507844

QY 564 ttctactcgtgtgaggggctgcgaccgcagctgcccgtggacatctacatccagg 623
Db 3507845 ---ctatgcgctcgtgcaggcggtgagatcgtgtgtccggctcgacatctacatccagg 3507901

QY 624 ctgcccacactacgcccagggccctgcctcctacggcagctcagctcagaggaagatcaa 683
Db 3507902 ctgcccgcgcgcgcggagatgctgtgcgcacgaatcctggaagctgcacgaaaaagattca 3507961

QY 684 gcggggagc 691
Db 3507962 gcagatgc 3507969

RESULT 4
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
```

```
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match      10.7%; Score 88; DB 4; Length 4411529;
Best Local Similarity 59.8%; Pred. No. 5.4e-11;
Matches 292; Conservative 0; Mismatches 190; Indels 6; Gaps 2;

QY 207 gctgcctcgcacaaaggccagagccgtggtcccaaccacccagcagccggcgagtgatgt 266
Db 3512049 gctgcagctgggattgaggtaggcgctgggactggagaaacagctgcccgccggatcct 3512108

QY 267 gctgcacagctgagatgacctcgtcaactggccggcgagttctctgtggcccatgac 326
Db 3512109 gctgtcacctgcagaaggctggcggtctatgtccgcaaaaactcctgtggccggcaac 3512168

QY 327 ctccggcctgcctgctgcgcgtggagatgacacatgacagcaccgcccgctacacat 386
Db 3512169 attcgattggcgtgctgtgcgcatcagatgctggcagccgggacaaagtttgacat 3512228

QY 387 ggaacgcgtttggcggtg---ttctccgcgcagccgcgcagctccgagctcagatcgt 443
Db 3512229 tgcgcgttccggtggaacggttctcgcgcagccgcgcagcgagcagatctgatcgt 3512288

QY 444 ggcgcgcacacacccaaagatggcccgagcttcgcgcagccgcgcagcgagcagatgcc 503
Db 3512289 ggcgcgcgcgcgcagcagagatggcgcgttactgcgcgcagatctatgacagatggc 3512348

QY 504 ggaacgcgcgtcagctgggtctccatgggagcgtgcgcacacggaggggctactaccacaa 563
Db 3512349 ggaacgcgaatgggttctggccatgggtgtgtgcgcctcgtcagtgaggattcaacaa 3512408

QY 564 ttctactcgtgtgaggggctgcgaccgcagctgcccgtggacatctacatccagg 623
Db 3512409 ---ctatgcgctcgtgcaggcggtgagatcgtgtgtccggctcgacatctacatccagg 3512465

QY 624 ctgcccacactacgcccagggccctgcctcctacggcagctcagctcagaggaagatcaa 683
Db 3512466 ctgcccgcgcgcgcggagatgctgtgcgcacgaatcctggaagctgcacgaaaaagattca 3512525

QY 684 gcggggagc 691
Db 3512526 gcagatgc 3512533

RESULT 5
5244792-2
; Patent No. 5244792
; APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPROTEIN
; B FROM HERPES SIMPLEX VIRUS
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,179
; FILING DATE: 20-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 921,730
; FILING DATE: 20-OCT-1986
; APPLICATION NUMBER: 597,784
; FILING DATE: 06-APR-1984
; SEQ ID NO: 2
; LENGTH: 3472
5244792-2
```

Query Match 5.6%; Score 46; DB 6; Length 3472;
 Best Local Similarity 43.9%; Pred. No. 0.049;
 Matches 196; Conservative 0; Mismatches 250; Indels 0; Gaps 0;

QY 247 agcagcggggcagatgtgtgtggccagagctggatgacatgtaacatggcccgccgg 306
 Db 1116 aactgcatactcagaggggtggagcgctcgtgtgtacccctacagagagttgtgtg 1175
 QY 307 agttctctgtggccatgacatctggccctggcctgtgtgcccgtggagatgatgacatg 366
 Db 1176 gcaactggcagactttgtatagtcccggtttacggctacagggaggggtctgcaaac 1235
 QY 367 gcagaccccgctacagatgagccgcttttggcgtgtgttcttcgagccagcccgccag 426
 Db 1236 gaacacaccagctacccgcccagccgtttcaagcagdggtgacgggtttctacgcgcgac 1295
 QY 427 tccgaactcatgatcgtggcgccgacacactcaacaaagatggcccagcgttcgcaag 486
 Db 1296 ctcaaccagggccggccagcggcgccagccaccccggaacctgctcaagaccaccag 1355
 QY 487 gtctacacacagatgcggcgccgctgactggtgtctccatggggagctggcccaacgga 546
 Db 1356 ttcacctggcctgggactgggtgcacaaagcccgctcgtgtgcacctgacccaagtgg 1415
 QY 547 ggaggtactaccactattctactcgtgtgtggagggctgacccgcatgtgcccgtg 606
 Db 1416 caggagtgacagatgtcgtctcagtgacggcgtctcttcctcgtattctctccac 1475
 QY 607 gacatctacatccagctgcccactacgagccgagccctgctctacggcatcctgag 666
 Db 1476 gccatctccacccttcacacacacacacacacacacacacacacacacacacacac 1535
 QY 667 ctgcagaggagatcaagcggagcg 692
 Db 1536 ggggactgcctcggaagcggcccg 1561

RESULT 6
 US-09-593-589-3
 ; Sequence 3, Application US/09593589
 ; Patent No. 6306655
 ; GENERAL INFORMATION:
 ; APPLICANT: Brett P. Monia
 ; APPLICANT: Madeline M. Butler
 ; APPLICANT: Jacqueline Wyatt
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP ALPHA EXPRESSION
 ; FILE REFERENCE: RTS-0119
 ; CURRENT APPLICATION NUMBER: US/09/593,589
 ; CURRENT FILING DATE: 2000-06-13
 ; NUMBER OF SEQ ID NOS: 94
 ; SEQ ID NO 3
 ; LENGTH: 3318
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (592)...(1658)
 US-09-593-589-3

Query Match 5.5%; Score 45.4; DB 4; Length 3318;
 Best Local Similarity 46.2%; Pred. No. 0.067;
 Matches 188; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

QY 326 ccttcggcctgctgctgctgctgagatgatgcacatggcagcaccgccctacgaca 385
 Db 680 ccttcggccttcccgggggcgcccgccgagcctcccgcccccacacgtgcgcccgg 739
 QY 386 tggacgccttttgcgtgttcttcggcgcagcccgccagctccgacatgatcgtgg 445
 Db 740 agccgctggcgcatctcgagacagacgacgtccatcgacatcagcgcctacatcgacc 799

QY 446 ccggcacactcaccaacaagatggccccagcgttctgcgaaggtctacgacacagatgcgg 505
 Db 800 cgccgcttcaacgacgagttctctggccgacctgttccagcacagcggcgacagagaga 859
 QY 506 agcc---gcctcagtggtctccatgggagctggcccaacggagggaggtactactacc 562
 Db 860 agcccaagggcggtgggcccccaagggcgggcgggcgccgactttgactaccgg 919
 QY 563 attcctactcgggtggaggggctgcgacgcgcatgtgcccgtggacatctacatcccaag 622
 Db 920 gcggccccggccccggcgcgccgtcatgccggggggagcgacgggcccccgccgg 979
 QY 623 gctgcccacactacggccggagccctgctctacggcatcctgcagctgcagagagatca 682
 Db 980 gctacggctgcggcgccggctacactggacggcaggtgtgagccctgtacgagcg 1039
 QY 683 agcggagggagggagctgcagatctgtgtacgcaggtagcgcgcgcg 729
 Db 1040 tcggggcgccggcgtgcggccgctggtgatcaagcagggagccccgc 1086

RESULT 7
 US-08-804-439A-13
 ; Sequence 13, Application US/08804439A
 ; Patent No. 6015565
 ; GENERAL INFORMATION:
 ; APPLICANT: Rose, Timothy M.
 ; APPLICANT: Bosch, Marnix L.
 ; APPLICANT: Strand, Kurt
 ; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHFV/KSHV
 ; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Ste 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,439A
 ; FILING DATE: February 21, 1997
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 09176/004001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 678-5070
 ; TELEFAX: (619) 678-5099
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2713 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 US-08-804-439A-13

Query Match 5.4%; Score 44.4; DB 3; Length 2713;
 Best Local Similarity 43.7%; Pred. No. 0.11;
 Matches 195; Conservative 0; Mismatches 251; Indels 0; Gaps 0;
 QY 247 agcagcggggcagatgtgtgtggccagcgtgagctgcctgcactgcacactggcccg 306
 Db 805 AACTGCATCGTCGAGGAGGTGACGCGCGCTCGGTGTACCGGTACGACGAGTTGTGTGCTG 864

Search completed: February 12, 2002, 18:09:48
Job time: 2851 sec

OM of: US-09-525-867-1 to: EST:* out_format : pfs

Date: Feb 12, 2002 7:52 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL-frame+2n.model -DEV-xlh
-q/cgn_1/USPRO_spool/US09525867/runat_12022002_111927_15597/app_query.fasta_1.272
-DB=EST -QFMT=fastap -SUFFIX=rat -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-OCAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-EGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09525867 -RCGNL_1_3758
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-525-867-1

Query length: 213

Database: EST*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 1179.910000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_est2:BG488658	+ 1108.00	1925.87	4.2e-98	695	! BG488658 602534588F1 NIH_MGC_18
gb_est2:BG334974	+ 1105.00	1920.51	8.3e-98	704	! BG334974 602403406F1 NIH_MGC_21
gb_est2:BG706832	+ 1102.00	1914.34	1.8e-97	774	! BG706832 602671984F1 NIH_MGC_96
gb_est2:BG479937	+ 1102.00	1913.60	2.0e-97	836	! BG479937 602527493F1 NIH_MGC_21
gb_est2:BG423354	+ 1080.00	1877.28	2.1e-95	804	! BF342354 602013145F1 NCI_CGAP_18
gb_est1:BE733056	+ 1080.00	1874.73	2.3e-95	711	! BE733056 601569557F1 NIH_MGC_21
gb_est2:BE794266	+ 1080.00	1874.67	3.0e-95	879	! BE794266 601591551F1 NIH_MGC_7
gb_est2:BG491103	+ 1068.00	1855.40	3.5e-94	738	! BG491103 602519937F1 NIH_MGC_18
gb_est2:BI116730	+ 1062.00	1844.98	1.3e-93	733	! BI116730 602868416F1 NIH_MGC_7
gb_est2:BI195871	+ 1059.50	1839.91	2.5e-93	788	! BI195871 602756115F1 NIH_MGC_19
gb_est1:AL580649	+ 1056.00	1833.66	5.7e-93	799	! AL580649 AL580649 LPI_NFL008_19
gb_est2:BE795373	+ 1054.00	1830.97	8.0e-93	736	! BE795373 601592903F1 NIH_MGC_7
gb_est2:BG419411	+ 1051.50	1824.54	1.8e-92	910	! BG419411 602446102F1 NIH_MGC_14
gb_est2:BG395342	+ 1051.00	1822.73	2.3e-92	1002	! BG395342 602457897F1 NIH_MGC_1
gb_est2:BG765790	+ 1049.00	1822.15	2.5e-92	742	! BG765790 602739620F1 NIH_MGC_49
gb_est2:BG206581	+ 1041.50	1808.65	1.4e-91	773	! BG206581 601870324F1 NIH_MGC_19
gb_est2:BG765478	+ 1031.00	1789.81	1.6e-90	813	! BG765478 602738979F1 NIH_MGC_49
gb_est2:BF968344	+ 1030.50	1787.57	2.1e-90	936	! BF968344 602694155F1 NIH_MGC_84
gb_est2:BI260297	+ 1029.00	1787.69	2.1e-90	706	! BI260297 602372896F1 NIH_MGC_12
gb_est2:BE897893	+ 1022.50	1774.16	1.2e-89	883	! BE897893 601440565F1 NIH_MGC_72
gb_est2:BI193338	+ 1012.00	1759.80	7.4e-89	586	! BI193338 602947659F1 NIH_MGC_42
gb_est2:BG336739	+ 992.00	1722.49	8.8e-87	747	! BG336739 602405173F1 NIH_MGC_21
gb_est2:BI133475	+ 984.00	1709.34	4.8e-86	686	! BI133475 602998099F1 NIH_MGC_12
gb_est2:BG031924	+ 947.00	1644.87	1.9e-82	574	! BG031924 602300616F1 NIH_MGC_87
gb_est2:BI260297	+ 936.50	1625.37	8.6e-82	717	! BI260297 601599428F1 NIH_MGC_21
gb_est1:AK003132	+ 930.50	1625.37	2.3e-81	759	! AK003132 Mus musculus adult mal
gb_est2:BF608952	+ 918.50	1593.83	1.3e-79	766	! BF608952 MY1_001894 Mouse 9-day
gb_est2:BG024362	+ 912.50	1584.40	4.3e-79	687	! BG024362 602747355F1 NIH_MGC_85
gb_est1:BE537419	+ 901.50	1564.70	5.4e-78	722	! BE537419 601058927F2 NIH_MGC_10
gb_est1:AL524830	+ 897.00	1554.42	2.0e-77	920	! AL524830 AL524830 LPI_NFL003_NE
gb_est2:BF486221	+ 893.50	1551.52	3.0e-77	665	! BF486221 UI-M-CO-ayf-a-04-0-0-UI
gb_est2:BI33809	+ 887.50	1542.50	9.4e-77	572	! BI33809 602999309F1 NIH_MGC_12
gb_est2:BE951803	+ 887.00	1539.63	1.4e-76	703	! BE951803 UI-M-CO-ayf-a-04-0-0-UI
gb_est2:BF119921	+ 878.00	1524.89	9.0e-76	695	! BF119921 601758003F1 NCI_CGAP_M
gb_est1:AJ393949	+ 878.00	1523.96	1.0e-75	699	! AJ393949 AJ393949 Gkfa226 Gallu
gb_est2:BG967059	+ 874.00	1514.77	3.3e-75	876	! BG967059 602834296F1 NCI_CGAP_0
gb_est1:AW245467	+ 867.00	1500.45	9.6e-75	586	! AW245467 28229001 prime NIH_MGC
gb_est2:BF079568	+ 861.00	1496.41	3.5e-74	560	! BF079568 202006 MARC 2P1G Sus s
gb_est2:BF671109	+ 832.00	1441.64	3.9e-71	854	! BF671109 602150934F1 NIH_MGC_81
gb_est2:BF682739	+ 831.50	1440.35	4.6e-71	891	! BF682739 602117105F1 Soares_mam

```
gb_est1:AW250024 - 821.00 1426.06 2.9e-70 587 ! AW250024 2819583.3prime NIH
gb_est2:BG683330 + 811.50 1405.95 3.8e-69 843 ! BG683330 602651438F1 NIH_M
gb_est1:AW546476 - 807.00 1401.35 6.8e-69 602 ! AW546476 1000805-3 NIA_N
gb_est2:BF682026 + 805.00 1396.03 1.4e-68 727 ! BF682026 602117105T1 Soare
gb_est1:BE563290 + 797.00 1381.32 8.9e-68 784 ! BE563290 60135414F1 NIH_M
```

seq_name: gb_est2:BG488658

seq_documentation_block:

```
LOCUS BG488658 695 bp mRNA EST
DEFINITION 602534588F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4662381.5',
mRNA sequence.
```

ACCESSION BG488658

VERSION BG488658.1 GI:13450165

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 695)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue procurement: DCTD/BTP/Gazdar

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

Plate: <http://image.llnl.gov>

High quality sequence stop: 691.

Location/Qualifiers

1..695

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4662381"

/tissue="NIH_MGC_18"

/lab_host="DH10B (phage-resistant)"

/note="origin: lung; Vector: pORF7; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

BASE COUNT 114 a 242 c 225 g 114 t

ORIGIN

alignment_scores:

Quality: 1108.00 Length: 213

Ratio: 5.226 Gaps: 0

Percent Similarity: 99.531 Percent Identity: 99.531

alignment_block:

US-09-525-867-1 x BG488658

Align seg 1/1 to: BG488658 from: 1 to: 595

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17

|||||

11 ATGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60

17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSer 34

|||||

61 GCCTCCAGCGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 110

34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50

```

|||||
111 TGGCCACCGATGGCCCAAGCAGCACCAGCGCTGCCCTGCCAAGGCCAGA 160
|||||
51 AlaValAlaProLysProSerArgGlyGluTyrValValAlaLysLe 67
|||||
161 GCCGTGGCTCCCAACCCAGCAGCGCGGCGAGTAGTGTGGTCCAGCT 210
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrp 84
|||||
211 GGATGACCTGCTCAACTGGGCGCCCGGAGTCTCTGTGGCCCATGACCT 260
|||||
84 heGlyLeuAlaCysCysAlaValAluMetMetHisMetAlaAlaProArg 100
|||||
261 TCGGCCCTGGCTGTGGCGCGTGAGATGATGCACATGGCAGCACCCTCC 310
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
311 TACGACATGGACCGCTTGTGGCTGTCTCCGCGCAGCGCGCCAGTC 360
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAla 134
|||||
361 CGAGCTCATGATCTGTGGCGGCACACTCACCACAGATGGCCCCAGCGC 410
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
411 TTCGAAGGCTACGACACAGATCGCGAGCGCGCTACGTGGTCTCCATG 460
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHistySerTyrSerValVa 167
|||||
461 GGGAGCTGCGGCAACGAGGAGGCTACTACCATATTCCTACGCTGTGT 510
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
511 GAGGGCTGCGACGCGCATGTCGCCGTGGACATCTACATCCAGGTCGCC 560
|||||
184 rOPrThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
561 CACCTACGGCGGAGGCGCTCTCTACGGCATCTGTCAGCTGCAGAGGAAG 610
|||||
201 lIleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
|||||
611 ATCAAGCGGAGCGGAGGCTGCAGATCTGTACCGCAGG 649

```

seq_name: gb_est2:BG334974

```

seq_documentation_block: 704 bp mRNA EST 27-FEB-2001
LOCUS BG334974 602403406F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4541304 5',
DEFINITION mRNA sequence.
ACCESSION BG334974
VERSION BG334974.1 GI:13141412
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 704)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999);
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1219 row: k column: 01
High quality sequence stop: 699.

```

FEATURES
source

1..704

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4541304"
/clone_lib="NIH_MGC_21"
/tissue_type="Choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 114 a 240 c 229 g 121 t

ORIGIN

alignment_scores:
Quality: 1105.00 Length: 214
Ratio: 5.188 Gaps: 1
Percent Similarity: 99.533 Percent Identity: 99.533

alignment_block:
US-09-525-867-1 x BG334974

Align seg 1/1 to: BG334974 from: 1 to: 704

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
4 ATGCGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 53
17 uArgSerSerValGlyProAlaValAlaGlnAlaArgGlyValHisGlnSerV 34
54 CGGCTCCAGCGTGGCGCGCGCTGTGCGGACGAGGTGTCATCAGAGCG 103
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
104 TGGCCACCGATGCCCCAAGCAGCAGCCAGCCGCTGCCCTGCCAAGGCCAGA 153
51 AlavaAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
154 CGCGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGGTGGCCAAAGCT 203
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrp 84
204 GGATGACCTGCTCAACTGGCGCGCGCGAGTCTCTGTGGCCCATGACCT 253
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
254 TCGGCTGGCTGCTGCTGCGCGTGGAGATGATGCACATGGCAGCACCCTCC 303
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
304 TAGCAGATGGACCGCTTGGCGTGGTCTTCGCGCCAGCGCGCCGCGCAGTC 353
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
354 CGACGTCATGATCGTGGCGGCGACACTCACCACAGATGGCCCGCCAGCGC 403
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
404 TTCGACAGGCTACGACCCAGATGCCGCGGCGCGGTACGTGGTCTCCATG 453
151 GlySerCysAlaAsnGlyGlyTyrTyrHistySerTyrSerValVa 167
454 GGGAGCTGCGGCAACGAGGAGGCTACTACCATATTCCTACGCTGTGTGT 503
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCys 183
504 GAGGGCTGGACCGCATGTCGCCGTGGACATCTACATCCAGGTCGC 553
184 ProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLy 200

|||||
 554 CCACCTACGGCGAGCCCTGCTACGGCATCCCTGCAGAGGAA 603
 200 sileLysArgGluArgLeuGlnIleTrpTyrArgArg 213
 |||||||
 604 GATCAAGCGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 643

seq_name: gb_est2:BG706832

seq_documentation_block:

LOCUS BG706832 774 bp mRNA EST 07-MAY-2001
 DEFINITION 602671984F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794669 5',
 mRNA sequence.

ACCESSION BG706832

VERSION BG706832.1 GI:13982569

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 774)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10676 row: g column: 22

High quality sequence stop: 762.

FEATURES

source

1..774

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4794669"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: Brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag

): oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',

size-selected for average insert size 2.3 kb and

normalized to R0T 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

132 a 259 c 251 g 132 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 1102.00

Ratio: 5.223 Length: 213

Percent Similarity: 99.061 Gaps: 0

Percent Identity: 99.061

alignment_block:

US-09-525-867-1 x BG706832 ..

Align seg 1/1 to: BG706832 from: 1 to: 774

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17

|||||

33 ATGCGGGTGTGACCTCTGCTGGCTGGCGGCTTCCGGATCCCTTGCT 82

|||||

17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnServ 34

|||||

83 ACCTCCACCGCTGGGCGCGCTGTGCAGCAGCAGGTGTCCATCAGACG 132
 34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
 |||||||
 133 TGCCACCGCATGTCGCCAAGCAGCAGCAGCAGCTGCTGCCAAAGGCCAGA 182
 51 AlaValAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
 |||||||
 183 GCGGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGGTGGCCAGCT 232
 67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
 |||||||
 233 GGATGACCTCGTCACTGGCGCGCGCGGAGTTCTCTGTGGCCCATGACCT 282
 84 heClyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
 |||||||
 283 TCGGCTGGCTGCTGCGCGCTGGAGATGATGCATGCAGCAGCAGCCAGC 332
 101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
 |||||||
 333 TAGGACATGGACCGCTTTGGCGTGTCTTCCGCGCCAGCCGCGCCAGTC 382
 117 rasPValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
 |||||||
 383 CGACGTCATGATCGTGGCGCGCACACTCACCACCAAGATGGCGCCAGCGC 432
 134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
 |||||||
 433 TTGCGAAGTCTACGACCAAGATGCGGAGCGCGCTACGTGTCTCCATG 482
 151 GlycerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
 |||||||
 483 GGGAGTGGCGCAACGAGGAGGCTACTACCACTATTCTACTCGGTGT 532
 167 largGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
 |||||||
 533 GAGGGCTCGACCGCATCGTCCCGTGACATCTACATCCAGCTTGC 582
 184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
 |||||||
 583 CACCTACGCGCCAGCGCTGTCTACGCGCATCTCGACGTCGAGAGGAAG 632
 201 IleLysArgGluArgLeuGlnIleTrpTyrArgArg 213
 |||||||
 633 ATCAACGGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 671

seq_name: gb_est2:BG479937

seq_documentation_block:

LOCUS BG479937 836 bp mRNA EST 21-MAR-2001

DEFINITION 603527493F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4651103 5',

mRNA sequence.

ACCESSION BG479937

VERSION BG479937.1 GI:13412216

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 836)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCMI433 row: i column: 24

FEATURES High quality sequence stop: 751.
Location/Qualifiers
source

1. .836
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4651103"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pORF7; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 152 a 281 c 261 g 141 t 1 others
ORIGIN

alignment_scores:
Quality: 1102.00 Length: 214
Ratio: 5.174 Gaps: 1
Percent Similarity: 99.533 Percent Identity: 99.065

alignment_block:

US-09-525-867-1 x BG479937 ..

Align seg 1/1 to: BG479937 from: 1 to: 836

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
7 ATGGGGTGTCTCAGCTCTGGGCTCGCGGCTCCGGATCTTGGTCT 56
|||||
17 uArgSerSer.ValGlyProAlaValGlnAlaArgGlyValHisGlnSer 33
|||||
57 GCGTCCAGCCCTGGGCGGCTGTGAGGACGAGGTGTCATCAGAGC 106
|||||
34 ValAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAla 50
|||||
107 GTGGCCACCGATGGCCCAAGCAGCACCGACCTGCTGCGCAAGGCCAG 156
|||||
50 GaLaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysL 67
|||||
157 AGCGGTGGCTCCCAACCCAGCAGCGGGCGAGATGTGTGGTGGCCAAAGC 206
|||||
67 euAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThr 83
|||||
207 TGGATGACCTGCTCACTGGGCGCGCGAGTCTCTGTGGCCCATGACC 256
|||||
84 PheGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProAr 100
|||||
257 TTCGGCCTGGCCTGTCGGCGGTGGAGATGATGCATGCGCAGCACCCCG 306
|||||
100 gTyRAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnS 117
|||||
307 CTACGACATGACCGCTTTGGCGTGTCTTCCGGCCAGCGCCCGCCAGT 356
|||||
117 erAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAla 133
|||||
357 CCGAGTGTATGTCGTGGCGGCACATCACCACCAAGATGGCCCGCAGCG 406
|||||
134 LeuArgLysValTyrAspGlnMetProGluProArgTyrValValSerMe 150
|||||
407 CTTGCAAGGCTCTACGACCATGCCGAGCGCGCTACGTGCTCCAT 456
|||||
150 tGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValV 167
|||||
457 GGGAGCTGCGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGGTG 506
|||||
167 aArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCys 183
|||||

507 TGAGGGCTGCGACCGCATCTGCCCTGGACATCTACATCCAGGCTGC 556
184 ProProThrAlaGluAlaLeuLeuTyrGlyLeuGlnLeuGlnArgTy 200
|||||
557 CCACCTACGGCCGAGGCCCTGCTCTACGGCATCTCGAGCTGCAGAGGA 606
|||||
200 sIleLysArgGluArgArgLeuGlnIleTrpTyrArgA-g 213
|||||
607 GATCAACGGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 646
|||||

seq_name: gb_est2:BF342354

seq documentation_block:
LOCUS BF342354 804 bp mRNA EST 22-NOV-2000
DEFINITION 602013145F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4148842
5', mRNA sequence.

ACCESSION BF342354
VERSION BF342354
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 804)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9409 row: j column: 11

High quality sequence start: 2

High quality sequence stop: 774.

FEATURES
source

Location/Qualifiers

1. 804

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCI_CGAP_Brn64"

/tissue_type="gliblastoma with EGFR amplification"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.57 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

268 c 255 g 132 t

149 a

BASE COUNT

ORIGIN

alignment_scores:

Quality: 1081.00 Length: 211

Ratio: 5.123 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-525-867-1 x BF342354

Align seg 1/1 to: BF342354 from: 1 to: 804

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
23 ATGGCGTGTGTGAGCTCTGGGCTCGCGGCTCCGGATCTTGGTCT 72
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
73 GCGTCCAGCGTGGCGGCGCTGTGAGGACGAGGTGTCATCAGAGCG 122
|||||

```

34  a1a1aThraspGlyProSerThrGlnProAlaLeuProLysAlaArg 50
|||||
123 TGCCACCACCGATGCCCAAGACACACCCAGCCCTGCCCTGCCAAAGGCCAGA 172
|||||
51  AlavalAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
|||||
173 GCGGTGGCTCCCAAAACCCAGCAGCGCGGAGTAGTGTGGTGGCCAGCT 222
|||||
67  uaspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
223 GGATGACCTCGTCACTGGCCCGCGGAGTCTCTGTGGCCCATGACCT 272
|||||
84  heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
273 TCGGCCTGGCTGCTGGCGCGGAGATGTCACATGGCAGCAGCCCGC 322
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
323 TAGCAGATGAGACCGCTTTGGCGTGTCTTCCGCGCAGCGCGCCAGTC 372
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
373 CGAGCTCATGATCGTGGCGCGGACACTCACACAAAGATGGCCCCAGCGC 422
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
423 TTGCGCAAGGTCTACGACCATGTCGGGAGCGCGCTAGTGTCTCCATG 472
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
473 GGAGTGTGGCGCAACGGAGAGGCTACTACCACTATTCTACTCGGTGGT 522
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
523 GAGGGCTCGCAGCGCATGTCGCCGCTGGACATCTACATCCAGGCTGCC 572
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuArgLys 200
|||||
573 CACCTACGCG. GAGGCGCTGCTCTACGCGATCTCTGACGCTCAGAGGA 621
|||||
201 IleLysArgGluArgArgLeuGlnIleTrpTyr 211
|||||
622 ATCAACGGGAGCGGAGGCTGCAGATCTGGTAC 654
|||||

```

seq_name: gb_est1.BE733056

```

seq_documentation_block:
LOCUS      BE733056          711 bp      mRNA          15-SEP-2000
DEFINITION 601569557F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844119 5',
            mRNA sequence.
ACCESSION  BE733056
VERSION    BE733056
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 711)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM540 row: a column: 16
High quality sequence stop: 711.

```

FEATURES

source
location/Qualifiers
1..711

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3844119"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      116 a      246 c      227 g      122 t
ORIGIN

```

alignment_scores:

Quality: 1080.00 Length: 205
Ratio: 5.268 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-525-867-1 x BE733056 ..

Align seg 1/1 to: BE733056 from: 1 to: 711

```

9  LeuArgGlyPheArgIleLeuGlyLeuArgSerValGlyProAlaVa 25
|||||
9  CTGCGCGGCTTCCGATCCTTGGTCTGCGCTCCAGCGTGGCGCGGTGT 58
|||||
25  lGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSer 42
|||||
59  GCAGGCACGAGGTGTCATCAGAGCGTGGCCACCGATGGCCCAACAGCA 108
|||||
42  hrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSerSer 58
|||||
109 CCGAGCTGCCCTGCCAAAGCGCAGAGCGTGGCTCCCAACCCAGCAGC 158
|||||
59  ArgGlyGlyTyrValValAlaLysLeuAspAspLeuValAsnTrpAlaAr 75
|||||
159 CGGGCGCATATGTGTGGCCAAAGCTGGATGACCTCGCACTCAACTGGGCCG 208
|||||
75  gArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValG 92
|||||
209 CCGGAGTTCTCTGTGGCCCATGACCTTCGGCTGGCCCTGCTGGCGCGTGG 258
|||||
92  lMetMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyVal 108
|||||
259 AGATGATGCATGGCAGCACCCCGCTACGACATGGCCGCTTTGGCGTG 308
|||||
109 ValPheArgAlaSerProArgGlnSerAspValMetIleValAlaGlyTh 125
|||||
309 GTCTTCGGCGCGCAGCCCGCGCAGTCCGACGTCATGATCGTGGCGCGCAC 358
|||||
125 rLeuThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetP 142
|||||
359 ACTCAACCAACAGATGGGCCCGCGCTTCGCAAGGCTTACGACCATGTC 408
|||||
142 roGluProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyGly 158
|||||
409 CGAGCGCGCGCTACGTGGTCTCCATGGGAGCTGGCCCAACCGAGGAGGC 458
|||||
159 TyrTyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValPr 175
|||||
459 TACTACCACTATCTCTACTCGGTGGTGGGGGCTGGACCGCATCGTGGC 508
|||||
175 oValAspIleTyrIleProGlyCysProThrAlaGluAlaLeuLeuT 192
|||||
509 CGTGGACATCTACATCCAGGCTGCCACCTACGCGCGGAGGCCCTGTCT 558
|||||
192 yrGlyIleLeuGlnLeuGlnArgLysIleLysArgGluArgLeuGln 208
|||||

```

```

|||||
559 ACGGATCCTCGAGTGCAGAGGAGATCAAGCGGAGCGGAGCGTGCAG 608
209 lletpTyArgArg 213
|||||
609 ATCTGGTACCGCAGG 623

```

seq_name: gb_est2:BE794266

```

seq_documentation_block:
LOCUS BE794266 879 bp mRNA EST 20-SEP-2000
DEFINITION 60159151F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3945589 5',
mRNA sequence.
ACCESSION BE794266
VERSION BE794266
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 879)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLM804 row: e column: 14
High quality sequence stop: 773.
Location/Qualifiers
1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3945589"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; site:1: XhoI; site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(S). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 166 a 276 c 280 g 157 t
ORIGIN
1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3945589"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; site:1: XhoI; site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(S). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

FEATURES

source

```

1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3945589"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; site:1: XhoI; site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(S). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 166 a 276 c 280 g 157 t

ORIGIN

alignment_scores:

```

Quality: 1080.00 Length: 213
Ratio: 5.118 Gaps: 0
Percent Similarity: 99.061 Percent Identity: 99.061
```

alignment_block:

US-09-525-867-1 x BE794266 ..

Align seg 1/1 to: BE794266 from: 1 to: 879

```

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleuLeuGlyLe 17
|||||
22 ATGGCGGGTGTGTGAGTCTGTGGCTGCGGGCTCCGATCTCGTGTCT 71
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
72 GCGTCCAGCGTGGGCTGTGCTGTGCAGGCACGAGGTGTCCATCAGAGCG 121
|||||

```

FEATURES

Location/Qualifiers

```

34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
122 TGGCCACCGATGGCCCAAGCAGCACCAGCTGCCCTGCCAAAGGCCAGA 171
|||||
51 AlaValAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
|||||
172 GCGGTGGCTCCAAACCCAGCAGCCGGGGGAGTATGTGTGGCCAAAGCT 221
|||||
57 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
222 GGATGACCTGTCAACTGGGCGCGCGGAGTTCTGTGGCCCATGACCT 271
|||||
84 heGlyLeuAlaCysCysAlaValGlnMetMetHisMetAlaAlaProArg 100
|||||
272 TCGGCTGGCTGTGGCGGTGGAGATGATGCACATGGCAGCACCCCGC 321
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
322 TACGACATGGACCGCTTGGCTGTTCGGCGCAGCCGCGCCAGTC 371
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
372 CGAGCTCATGATGTGGCGCGCACACTCACCAACAGATGGCCCCAGCGC 421
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
422 TTCGCAAGGTCTACGACCATGCGGAGCGCGCTACGTGTCTCCATG 471
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
472 GGGAGCTGGCCCAACGAGGAGGACTACTACCTATTCCTACTCGGTGGT 521
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
522 GAGGGGTGGCAGCG.ATCGTGGCGGTGGACATCTACATCCAGGCTGCC 570
|||||
184 roPcThrAlaGluAlaLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
571 CACCTAGCGCGGAGGCGCTCTAGGCGATCTGTGAGTGTGCAGAGGAAG 620
|||||
621 ATCAAGCGGAGGAGGAGGCTGCAGATCTGTGTACCGCAGG 659
|||||
seq_name: gb_est2:BG491103
seq_documentation_block:
LOCUS BG491103 738 bp mRNA EST 27-MAR-2001
DEFINITION 602519337F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4638441 5',
mRNA sequence.
ACCESSION BG491103
VERSION BG491103.1 GI:13452615
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 738)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1400 row: j column: 10
High quality sequence stop: 738.
Location/Qualifiers

```

```

source
1. 738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4638441"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 125 a 252 c 235 g 126 t
ORIGIN

alignment_scores:
Quality: 1068.00 Length: 211
Ratio: 5.110 Gaps: 0
Percent Similarity: 99.052 Percent Identity: 99.052

alignment_block:
US-09-525-867-1 x BG491103
Align seg 1/1 to: BG491103 from: 1 to: 738

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
7 ATGGCGTGTCTCAGCTCTCTGGCTGGCGGCTTCCGGATCCTTGGTCT 56
|||||
17 uArgSerValGlyProAlaValAlaAargGlyValHisGlnServ 34
|||||
57 GCCTCCAGCTGGGCTGGCTGTGCAGCAGAGGTGTTCATCAGACG 106
|||||
34 alaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
107 TGGCCACCGATGCCCAAGCAGCAGCCAGCCTGCCCTGCCAAGSCCAGA 156
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyValValAlaLysLe 67
|||||
157 GCGTGGCTCCCAACCCAGCAGCGGGGCGAGTATGGTGGCCAAAGCT 206
|||||
67 uAspAspLeuValAsnTrpAlaAargSerSerLeuTrpProMetThr 84
|||||
207 GGATGACCTCGTCACTGGCCCGCGGAGTCTCTGTGGCCCATGACCT 256
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetalAlaProArg 100
|||||
257 TCGGCTGGCTGCTGCTGCGCGTGGAGATGATGCATGCGCAGCAGCCGCG 306
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
307 TAGGACATGACCGCTTTGGCTGGTCTTCCGCGCAGCCCGCGCCAGTC 356
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetalAlaProAla 134
|||||
357 CGACGTCATGATCGTGGCGGCACACTACCAACAAGATGGCCCGCAGCGC 406
|||||
134 euArgLysValTyAspGlnMetProGluProArgTyValValSerMet 150
|||||
407 TTCGCAAGGCTTACGACCAAGATGCCGAGCCGCGTACGTGGTCTCCATG 456
|||||
151 GlySerCysAlaAsnGlyGlyTyTrpHisTySerTySerValVa 167
|||||
457 GGGAGCTGGCCCAACGAGAGGCTACTACCACTATTCCTACTCGTGTGT 506
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrlleProGlyCysP 184
|||||
507 GAGGGCTGGCAGCGATGTCGCCGTGGACATCTACATCCCGAGCTGCC 555

```

```

184 roProThrAlaGluAlaLeuLeuTyGlyIleLeuGlnLeuGlnArgLys 200
|||||
556 A.CCTACGCCGAGGCGCTCTACGGCATCTCGAGCTGCAGAGGAAG 605
|||||
201 IleLysArgGluArgGluGlnIleTrpTyr 211
|||||
606 ATCAACGGGAGCGGAGGCTGCAGATCTGGTAC 638
|||||
seq_name: gb_est2:B116730

seq_documentation_block:
LOCUS B116730 733 bp mRNA EST 26-JUN-2001
DEFINITION 602868418F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5017179 5',
mRNA sequence.
ACCESSION B116730
VERSION B116730.1 GI:14567631
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1828 row: C column: 04
High quality sequence stop: 670.
Location/Qualifiers
1. 733
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5017179"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 130 a 253 c 230 g 120 t
ORIGIN
```

```

alignment_scores:
Quality: 1062.00 Length: 204
Ratio: 5.232 Gaps: 0
Percent Similarity: 99.510 Percent Identity: 99.020
```

```

alignment_block:
US-09-525-867-1 x B116730
Align seg 1/1 to: B116730 from: 1 to: 733

10 ArgGlyPheArgIleLeuGlyLeuArgSerValGlyProAlaValG 26
|||||
3 CGCGGCTTCGGATCCTTGGTCTGGCTCCAGCTGGCGCTGGCTGTGCA 52
|||||
26 nAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSerThrG 43
|||||
53 GGCACGAGGTGTCATCATCAGAGGTGGCCACCATGGCCCAAGCAGCACCC 102
|||||

```

```

43  InProAlaLeuProLysAlaArgAlaValAlaProLysProSerSerArg 59
|||||
103 AGCCTGCCCTGCCAAGCCAGAGCCGTGGCTCCCAACCCAGCAGCCGG 152
|||||
60 GlyGluTyrValValAlaLysLeuAspLeuValAlaSerTrpAlaArgAr 76
|||||
153 GCGAGTATGTGGTGGCAAGCTGGATGACCTGCTCAACTGGGCCCGCG 202
|||||
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluM 93
|||||
203 GAGTCTCTGTGGCCCATGACCTTGGGCTGGCCCTGCTGGCCGTGAGA 252
|||||
93 etMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
|||||
253 TGATGCACATGACGACCCCGCTAGGACATGGACCGCTTTGGGCTGTC 302
|||||
110 PheArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLe 126
|||||
303 TTCCGCGCCAGCCCGCCGACGTCGACGTCATGATCGTGGCCGACACT 352
|||||
126 uThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
|||||
353 CACCAACAGATGGCCCGCAGCGCTTCGCAAGCTCTAGCAGCATGCCGG 402
|||||
143 luProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
|||||
403 AGCCGCGCTACGTGCTCTCCATGGGAGCTGCGCCAAACGGAGGCTAC 452
|||||
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176
|||||
453 TACCACTATCTCTACTCGTGGTGAGGGGCTGCGACCGCATCGTGCCT 502
|||||
176 lAspIleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG 193
|||||
503 GGACATCATACATCCAGCGCTGCCACCTACGTGGAGGCCCTGCTCTACG 552
|||||
193 lYleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIle 209
|||||
553 GCATCTCAGCTGCAGAGGAAGATCAAGCGGGAGCGGAGGCTGCAGATC 602
|||||
210 TrpTyrArgArg 213
|||||
603 TGGTACCGCAGG 614

seq_name: gb_est2:BI195871

seq_documentation_block:
LOCUS BI195871 788 bp mRNA EST 10-JUL-2001
DEFINITION 602756115F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4891946 5',
mRNA sequence.
ACCESSION BI195871
VERSION BI195871.1 GI:14650891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 788)
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1771 row: m column: 03
High quality sequence stop: 770.
Location/Qualifiers
FEATURES

```

source 1. 788

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4891946"

/clone_lib="NIH_MGC_19"

/tissue_type="neuroblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 147 a 254 c 254 g 133 t

ORIGIN

alignment_scores:

Quality: 1059.50 Length: 212

Ratio: 5.094 Gaps: 2

Percent Similarity: 98.113 Percent Identity: 97.642

alignment_block:

US-09-525-867-1 x BI195871

Align seg 1/1 to: BI195871 from: 1 to: 788

3 ValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSe 19

29 GTGCTGTGACGCTCTGGCTGCGGGGCTTCCGGATCCTTGTGTCGCTC 78

19 rSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerValAla 36

79 CAGCGTGGCCCGCGCTGTGACGACACCGAGGTGTCCATCAGAGCGTGG 128

36 hrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVal 52

129 CCGATGGCCCAAGCAGCAGCAGCCAGCGCTGCCCTGCCAAAGCCAGAG 178

53 AlaProLysProSerSerArgGlyGluTyrValValAlaLysLeuAsp 69

179 GCTCCCAACCCAGCAGCGGGGCGAGTATGTGTGCCAAGCTGGATGA 228

69 pLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrPheGly 86

229 CCGTCAACTGGCCCGCGCGAGTTCCTGTGGCCCATGACCTTCGGCC 278

86 euAlaCysCysAlaValAlaGluMetMetHisMetAlaAlaProArgTyr 102

279 TGGCGTGTGCGCGGTGGAGATGATGCACATGGCAGACCCCGCTACGAC 328

103 MetAspArgPheGlyValValPheArgAlaSerProArgGlnSerAsp 119

329 ATGGACCGCTTTGGCGTGGTCTCCGCGCCAGCCGCGCCAGTCCGACGT 378

119 lMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgL 136

379 CATGATCGTGGCCCGCACACTCACCAACAGATGGCCCGCCAGCGCTTC 428

136 ysValTyrAspGlnMetProGluProArgTyrValValSerMetGlySer 152

429 AGGTCTAGCACCAGATGCGCGAGCCGCGCTACGTGGTCTCCATGGGGAG 478

153 CysAlaAsnGlyGlyGlyTyrTyrHisTyrSerTyrSerValValArgG 169

479 TGCGCCAACGGAGAGGCTACTACCACTATTCCTACTCGGTGGTGGGG 528

169 yCysAspArgIleValProValAspIleTyrIleProGlyCysProPro 185

529 CTGCGACCGCATCGTCCCGTGGACATCTACATCCCGAGGCTGCCACC... 575


```

/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes: Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
BASE COUNT      127 a      244 c      240 g      125 t
ORIGIN

alignment_scores:
  Quality: 1054.00      Length:      213
  Ratio: 4.995          Gaps: 0
  Percent Similarity: 99.061      Percent Identity: 99.061

alignment_block:
US-09-525-867-1 x BE795373

Align seg 1/1 to: BE795373 from: 1 to: 736

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
8 ATGGGGGTGCTGTACGCTCCGCTGGCTG.CGCGGCTTCGGATCCTGGGTCT 56
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
57 GCGCTCCAGCTGGCGCTGGCTGTGCAGGACGAGGTGTCCATCAGAGCG 106
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
107 TGGCCACCGATGGCCCAAGCAGCACCAGCTGCCCTGCCAAGGCCAGA 156
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
157 GCGGTGGCTCCCAACCCAGCAGCGGGGGAGTATGTGTGGCCAAAGCT 206
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
207 GGATGACCTGCTCAACTGGGCGCCCGGAGTTCTCTGTGGCCCATGACCT 256
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
257 TGGGCTGGCTGTGCGCGCTGGAGATGATGCACATGGCAGCACCCCGC 306
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
307 TACGACATGGACCGCTTGGCGTGGTCTTCGCGCCAGCCCGCCGATGC 356
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
357 CGACGTATGATGCTGGCCGCGCACACACACCAAGATGGCCCGCCAGCG 406
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
407 TTCGCAAGTCTAGACCAAGTGGCGGAGCGCGCTACGTGTCTCCATG 456
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVal 167
|||||
457 GGGAGCTGGCCCAACGGAGGAGGCTACTACCACTATTCTCTACTCGGTGT 506
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
507 GAGGGCTGGCCAGCATCGTGGCCCGGGACATCTACATCCCGAGGTGCC 556
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuAlaArgLys 200
|||||
557 CACCTACGGC.GAGGCCCTGCTCTACGGCATCTCGAGCTGCAGAGGGAG 605
|||||

```

```

201 IleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
|||||
606 ATCAAGCGGAGCGAGGAGCTGCAGATCTGGTACGCGAGG 644

```

```
seq_name: gb_est2:BG419411
```

```

seq_documentation_block:
LOCUS      BG419411          910 bp      mRNA      EST      14-MAR-2001
DEFINITION 60246102F1 NTH_MGC_14 Homo sapiens cDNA clone IMAGE:4584528 5',
trna sequence.
ACCESSION  BG419411
VERSION    BG419411.1  GI:13325917
KEYWORDS   EST
SOURCE     human.
ORGANISM   Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 910)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: DCTD/DTF
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCMI310 row: d column: 01
          High quality sequence stop: 702.

```

```

FEATURES
    source
        1..910
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4584528"
            /tissue_type="renal cell adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      195 a      279 c      278 g      156 t
ORIGIN

alignment_scores:
  Quality: 1051.50      Length:      211
  Ratio: 5.104          Gaps: 1
  Percent Similarity: 97.630      Percent Identity: 96.682

```

```
alignment_block:
US-09-525-867-1 x BG419411
```

```
Align seg 1/1 to: BG419411 from: 1 to: 910
```

```

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
8 ATGGCGGTGCTGTACGCTCGCTGGCTGGCGGGTTCGGGATCCTGGTCT 57
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
58 GCGCTCCAGCGTGGCGCTGGCTGTGCAGGACGAGGTGTCCATCAGAGCG 107
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
108 TGGCCACCGATGGCCCAAGCAGCACCAGCTGCCCTGCCAAGGCCAGA 157
|||||

```

```

51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
158 GCCGTGCTCCAAACCCAGCACCGCGGAGTAGTGTGTGGCCAAAGCT 207
|||||
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
208 GGATGACCTCGTCAACTGGCCCGCGGAGTCTCTGTGGCCCATGACCT 257
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
258 TCGGCCTGGGCTCTGCGCGTGGAGATGATGCACATGGCAGCACCCCGC 307
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
308 TACACATGACACCGCTTGGCGTGGTCTTCCGGCCGAGCCCGCCAGTTC 357
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
358 CGAGCTCATGATCGTGGCCGGCACACTCACCAACAAGATGGCCCGAGCC 407
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
408 TTCACAGGTCTACGACACAGATCCGAGCCGCGTACGTGTCTCCATG 457
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
458 GGGAGCTGGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGTGTGT 507
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
508 GAGGGGTGGGACCGCATCGTGGCCGTGGACATCTACATCCCGAGGTGC. 556
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
557 ..CACCTACCGCGAGGCGCTCTCTACGGCATCTCGAGCTGCAGAGGAAG 604
|||||
201 lIleLysArgGluArgArgLeuGlnIleTrpTyr 211
|||||
605 ATCAAGCGGAGCGAGGCTGCAGATCTGTGATC 637
|||||

seq_name: gb_est2:BG395342

seq_documentation_block:
LOCUS BG395342 1002 bp mRNA EST 12-MAR-2001
DEFINITION 6024578791 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580538 5',
mRNA sequence.
ACCESSION BG395342
VERSION BG395342.1 GI:13288790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1299 row: m column: 19
High quality sequence stop: 753.
Location/Qualifiers
1..1002
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4580538"

```

```

/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

```

BASE COUNT      215 a      313 c      321 g      153 t
ORIGIN

```

```

alignment_scores:
  Quality: 1051.00      Length: 216
  Ratio: 5.005          Gaps: 3
  Percent Similarity: 97.222      Percent Identity: 97.222

```

```

alignment_block:

```

```

US-09-525-867-1 x BG395342

```

```

Align seg 1/1 to: BG395342 from: 1 to: 1002

```

```

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
20 ATGGCGTGTGTTCAGCTCTCTGGCTTCGGCGGCTTCCGGATCCTTGGTCT 69
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
70 GCGTCCAGCGTGGGCTGTGCTGTGCAGGCAGCGAGGTGTCCATCAGAGCG 119
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
120 TGGCACCGATGGCCCAAGCAGCACCCAGCTGCTGCCAAAGGCCAGA 169
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
170 GCGTGTGCTCCCAACCCAGCACCGCGGCGAGTAGTGTGGCCAAAGCT 219
|||||
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
220 GGATGATCTCGTCAACTGGCGCCGCGGAGTCTCTGTGGCCCATGACCT 269
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
270 TCGGCTGGGCTCTGCTGGCGGTGGAGATGATGCACATGGCAGCACCCCGC 319
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
320 TAGCAGATGGACCGCTTGGCGTGTCTTCCGCGCCAGCCCGCCAGTTC 369
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
370 CGAGCTCATGATCGTGGCGGACACTCACCAACAAGATGGCCCGCCAGCC 419
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
420 TTCGACAGGTCTACGACACAGATCCCGGAGCCGCGCTACGTGTCTCCATG 469
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
470 GGGAGTGGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGGTGTGT 519
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
520 GAGGGGTGGGACCGCATCGTGGCGGTGGACATCTACATCCCAAGGCTGC 569
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArg 199
|||||
570 CACCTACGCGCGAGGCGCTTGTCTACGGCATTCCTGCAGGTTCAGAGG 619
|||||

```


200 LysileuArgGluArgGlnLeuGlnIleTyrArgArg 213
 |||
 620 AAGTACAGCGGAGGAGGAGGCTGCAGATCTGGTACCGCAGG 661
 seq_name: gb_est2:BG765790

seq_documentation_block: 742 bp mRNA PST 15-MAY-2001
 LOCUS BG765790 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4869532 5',
 DEFINITION 602739620F1
 mRNA sequence.

ACCESSION BG765790
 VERSION BG765790.1 GI:14076443
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 742)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LICM1741 row: g column: 05
 High quality sequence stop: 739.
 Location/Qualifiers

FEATURES

1..742
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IWAG:4869532"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library..!"
 BASE COUNT 155 a 245 c 224 g 118 t
 ORIGIN

alignment_scores:

Quality: 1049.00 Length: 201
 Ratio: 5.245 Gaps: 0
 Percent Similarity: 99.502 Percent Identity: 99.502

alignment_block:

US-09-525-867-1 x BG765790 ..

Align seg 1/1 to: BG765790 from: 1 to: 742

13 ArgIleuGlyLeuArgSerValGlyProAlaValGlnAlaArgG1 29
 |||
 3 CGGATCCTTGGTCTCGCTCCAGCGTGGGCTGGCTGTGCAGGACGAGG 52
 |||
 29 yValHlsGlnSerValAlaThrAspGlyProSerSerThrGlnProAla 46
 |||
 53 TGTCCATCAGAGCGTGGCAGCCAGATGGCCCAAGCAGCAGCCAGCGCTGCC 102
 |||
 46 euProLysAlaArgAlaValAlaProLysProSerSerArgGlyGluTyr 62
 |||

103 TCCCAAGGCCAGAGCCGCTGGCTCCCAAAACCCAGCAGCCGGGGGAGTAT 152
 |||
 63 ValValAlaLysLeuAspAspLeuValAsnTrpAlaArgArgSerSerLe 79
 |||
 153 GTGGTGGCCCAAGCTGGATGACCTCGTCAACTGGGCCCGCGAGTTCTCT 202
 |||
 79 utrPrMotThrPheGlyLeuAlaCysCysAlaValGluMetMethHism 96
 |||
 203 GTGGGCCCATGACCTTCGGCCTGGCTGCTGGCGGTGGAGATGATGCACA 252
 |||
 96 etAlaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAla 112
 |||
 253 TGGCAGCACCCCGCTACGACATGGACCGCTTTGGCGTGGTCTCCGCGCC 302
 |||
 113 SerProArgGlnSerAspValMetIleValAlaGlyThrLeuThrAsnLy 129
 |||
 303 AGCCCGCGCCAGTCCGAGCTCATGATCGTGGCGGCACACTCACCAACAA 352
 |||
 129 sMetAlaProAlaLeuArgLysValTyrAspGlnMetProGluProArgT 146
 |||
 353 GATGCCCCAGCGCTTCGCAAGGTCTACGACCAGATGCCGGAGCGCGCT 402
 |||
 146, yzValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyr 162
 |||
 403 AGTGTGTCTCCATGGGAGCTCGCCCAACGGAGGAGGCTACTACCATAT 452
 |||
 163 SerTyrSerValValArgGlyCysAspArgIleValProValAspIleTy 179
 |||
 453 TCCTACTCGTGTGGTGGGCTGGCAGCGCATCGTCCCGTGGACATCTA 502
 |||
 179 rIleProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuG 196
 |||
 503 CATCCCAAGGCTGCCACCTACGGCCGAGGCGCTGCTCTACGGCATCTGC 552
 |||
 196 InLeuGlnArgLysIleLysArgGluArgLeuGlnIleTyrTyrArg 212
 |||
 553 ACCTCAGAGGAGAGATCAAGCGGGAGCGGAGCGCTGCAGATCTGTACCGC 602
 |||
 213 Arg 213
 |||
 603 AGG 605


```
564 TTTGGGGCCNCAAGAACTTTTGGGTGNCCTGNTGNGCCCGGAGGTGA 515
94 etHisMetAlaAa.ProArgTyArgMetAspArg.PheGlyValValP 110
110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126
464 TTTGGGCCACCCCGCCAGTCTGAGTNAATGATTGTGGCNGGNANANTA 415
127 ThrAsnLysMetAlaProAlaLeuArgLysValTyArgMetProG1 143
414 ACCAAAAGAGGGCCCGCCAGNCTTGC.AAGGTTTAAGACACAGATGCCGA 366
143 uprArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyT 160
365 CCCCGANAAAGTGTTCACGAGGAGTGN.GCCAAACGGNGGGTAAAN 317
160 yrHisTySerTySerValValArgGlyCysAspArgIleValProVal 176
316 ACCACTATTCCTAATNGTGTGAGGGGNTGCCACCGCATTTGCCCGTG 267
177 AspIleTyIleProGlyCysProProThrAlaGluAlaLeuLeuTyG1 193
266 GAATTTNAAATCCAGGGTCCACACNANGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
216 CATCCTGCAGTGCAGAGGAATCAAGCGGAGCGAGGCTGCANATT 167
210 rPyTyArgArg 213
166 GGTACCCGAGG 156
```

seq_name: /cgn2_6/prodata/2/ina/6B_COMB.seq:us-09-030-607-32

```
seq_documentation_block:
; Sequence 32, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030.607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: CDNA
US-09-030-607-32

alignment_scores:
  Quality: 411.00      Length: 189
  Ratio: 3.067        Gaps: 5
  Percent Similarity: 70.899  Percent Identity: 58.201

alignment_block:
US-09-525-867-1 x US-09-030-607-32/rev ..
Align seg 1/1 to reverse of: US-09-030-607-32 from: 1 to: 789

36 ThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVa 52
696 ACCGGGGGGCCCGAAGGGGAGCCANCTTCCNNCCCAAGNAGAGCGGTG 647
52 lAlaProLysProSerSerArgGlyGluTyValValAlaLysLeuAspA 69
646 GGITCCCAAAACCCAGNAAACGGGGGAGGATTGGGGGCCCAAGGGGATG 597
69 sp.LeuValAsnTrpAlaArgSerSerLeuTrpPro..... 81
596 ACITTTGTG.....AAACTGGGNCCTCCCGGAGNTTT 565
82 .....MetThrPheGlyLeuAlaCysCysAlaValGluMetM 94
564 TTTGNGGGCCNCAAGAACTTTTGGGTGNCCTGNTGNGCCCGGAGGTGA 515
94 etHisMetAlaAa.ProArgTyArgMetAspArg.PheGlyValValP 110
514 TGNANATNGCAGCAACCCCGNNAAGACATGNCCTGTTTGGGNGNTTT 465
110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126
464 TTTGGGCCACCCCGCCAGTCTGAGTNAATGATTGTGGCNGGNANANTA 415
127 ThrAsnLysMetAlaProAlaLeuArgLysValTyArgMetProG1 143
414 ACCAAAAGAGGGCCCGCCAGNCTTGC.AAGGTTTAAGACACAGATGCCGA 366
143 uprArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyT 160
365 CCCCGANAAAGTGTTCACGAGGAGTGN.GCCAAACGGNGGGTAAAN 317
160 yrHisTySerTySerValValArgGlyCysAspArgIleValProVal 176
316 ACCACTATTCCTAATNGTGTGAGGGGNTGCCACCGCATTTGCCCGTG 267
177 AspIleTyIleProGlyCysProProThrAlaGluAlaLeuLeuTyG1 193
266 GAATTTNAAATCCAGGGTCCACACNANGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
216 CATCCTGCAGCTGCAGAGGAANATCAAGCGGAGCGAGGCTGCANATT 167
210 rPyTyArgArg 213
166 GGTACCCGAGG 156
```

seq_name: /cgn2_6/prodata/2/ina/6B_COMB.seq:us-09-103-840A-2

```
seq_documentation_block:
; Sequence 2, Application US/09103840A
; Patent No. 6294928
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
```

```

; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

alignment_scores:
    Quality: 449.00      Length: 220
    Ratio: 3.054         Gaps: 6
    Percent Similarity: 66.818      Percent Identity: 41.818

alignment_block:
    US-09-525-867-1 x US-09-103-840A-2 ..
Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765
4 LeuSerAlaProGlyLeuArgGlyPheArgIle..... 14
3507435ATTTCATGCTCAGCGGTGTCTGGCCCTACGGGTATGTGGCGCCGGGGG 3507484
15 .....LeuGlyLeuArgSerValGlyProAlaValGlnAlaArgG 29
3507485GCCTGACGCTGGGATTGAGTAGGCGGTGGGA..... 3507515
29 lyValHisGlnSerValAlaThrAspGlyProSerSerThrGlnProAla 45
3507516.....CTGGAAGAACAG 3507527
46 LeuProLysAlaArgAlaValAlaProLysProSerSerArgGlyGluP 62
3507528CTGCCC.....GGCGGAT 3507541
62 rValValAlaLysLeuAspSerLeuValAsnTrpAlaArgSerSerL 79
3507542CCTCTGCTACCGCTCGAGAAGTGGCGGCTATGTCGCCAAAACTCCC 3507591
79 eutPpMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHis 95
3507592TGTGCGCGGCAACATCCGATTGCGGTGTCGCGATCGAGATGATGGCG 3507641
96 MetAlaAlaProArgTyAspMetAspArgPheGlyValVal...PheAr 111
3507642ACCCCGGGACCAAGTTTGACATTGCGCGGTTCCGGATGGAACGGTTC 3507691
111 gAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeuThra 128
3507692GGCCCGCGGACGACATCTGATGATCGTGGCGGCGCGGTGACGC 3507741
128 snLysMetAlaProAlaLeuArgLysValTyAspGlnMetProGluPro 144
3507742AGAAGATGGCGGTACTCGCCAGATCTATGACCGATGCGGAGCGG 3507791
145 ArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyTyThi 161
3507792AAATGCGTTCTGGCCATGGGTGTGTGGCTCTGTCAGTGGATGTCAA 3507841
161 sTySerSerValValArgGlyCysAspArgIleValProValAspi 178
3507842C...AACTATGCGATCTGCGAGCGGTGGATCATGTGTTCGGTCCACA 3507888
178 leTyIleProGlyCysProProThraAlaGluAlaLeuLeuTyGlyIle 194

```

```
111 gAlaserProArgGlnSerAspValMetIleValAlaGlyThrLeuThrA 128
|||||
3512256GGCCACGCGCGGAGCAGATCGATGATCGTGGCGGCGGTGACGC 3512305

128 snLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProGluPro 144
|||||
3512306AAGATGCGCGCGTACTGCGCGCAGATCATGACAGATGGCGGAGCG 3512355

145 ArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHI 161
|||||
3512356AAATGGGTCTGGCAGTGGGTCTGGCCTCGTCAGTGGGATGTCAA 3512405

161 sTySerTyrSerValValArgGlyCysAspArgIleValProValAspi 178
|||||
3512406C...AATATCGATCGTGCAGGCGTGGATCATGTGTTCGGTGCACA 3512452

178 leTyrIleProGlyCysProProThrAlaGluAlaLeuTyrGlyIle 194
|||||
3512453CTACCTACCGCGGTGCGCGCGCGCGCGAGATGCTGCTGCACGAATC 3512502

195 LeuGlnLeuGlnArgLysIleLys.....ArgGI 204
|||||
3512503CTGAAGCTGCAGAAAGATTCACAGATGCCATTAGGTATCAACCGGA 3512552

204 uArgArgLeu 207
|||||
3512553ACGCGCTATC 3512562
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-801-344-7

```
seq_documentation_block:
; Sequence 7, Application US/08801344
; Patent No. 6087140
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Alkarak, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
; TITLE OF INVENTION: FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.344
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3524 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pyridine nucleotide transhydrogenase gene
```

US-08-801-344-7

alignment_scores:

Quality: 101.00 Length: 218
Ratio: 0.971 Gaps: 11
Percent Similarity: 47.706 Percent Identity: 22.477

alignment_block:

US-09-525-867-1 x US-08-801-344-7 ..

Align seg 1/1 to: US-08-801-344-7 from: 1 to: 3524

```
8 GlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAl 24
|||||
1046 GGTGTTGAGTCTGGCGCGCATTTGGCGAGCAACAGTCTCGCGCGCAT 1095

24 aValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSer 41
|||||
1096 TGTG..... 1099

41 erThrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSer 57
|||||
1100 .....CGTGCATTCGACACCCCGCGGAA 1123

58 Ser ArgGlyGluTyrValValAlaLysLeuAspAspLeuValAsnTrpA 74
|||||
1124 GTGAAGACACAAGTTCAAAGTA.....TGGG 1149

74 laArgArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysCys... 89
|||||
1150 CGCGAATTCCTCGAGCTGATTATAAGAGGAAGTGGCAGCGCGCATG 1199

90 AlaVal..GluMetMetHisMetAlaAlaProArgTyrAspMetAspArg 105
|||||
1200 GCTATGCCAAAGTGTATCGGACGCGTTTCATCAAGCGGAAATGGAA... 1246

106 PheGlyValValPheArgAlaSerProArgGlnSerAspValMetIleVa 122
|||||
1247 .....CTCTTTGCCGCCAGGCAAAAGAGGTGATATCATTTGTCAC 1287

122 laAlaGlyThrLeuThrAsnLysMetAlaProAlaLeu.....ArgLysV 137
|||||
1288 CACCGCGCTATTCCAGGCAACACCGCGCGAAGCTAATTACCGGTGAAA 1337

137 alTyrAspGlnMetProGluProArgTyrValValSerMetGlySerCys 153
|||||
1338 TGGTTGACTCCATGAAGCGGCGCAGTGTGATTGTGCACCTG...GCAGCC 1384

154 AlaAsnGlyGlyGlyTyrTyrHisTyrSerTyrSerValValArgGlyCy 170
|||||
1385 CAAAACGGCGGC.....AACTG 1401

170 sAspArgIleValProValAspIleTyr..... 179
|||||
1402 TGAATACACCGTCCCGGTGAAATCTTCATCGAGAAATGGTGTCAAAG 1451

180 .....IleProGlyCysProProThrAlaGluAlaLeu 190
|||||
1452 TGATTGTTTATACCGATCTTCGCGGCGCTGCGCAGCAATCCTCACAG 1501

191 LeuTyrGly.....IleLeuGlnLeuGlnArgLysIleLysArgGluAr 205
|||||
1502 CTTTACGGCACAAACCTCGTTATCTGCTGAACACTGTTGTGCAAGAGAA 1551

205 g 205
1552 A 1552
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-498-599-7

seq_documentation_block:
; Sequence 7, Application US/09498599

```
; Patent No. 6303352
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF
; 1,2-PROPANEDIOL FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,599
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3524 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pyridine nucleotide transhydrogenase gene
; US-09-498-599-7

alignment_scores:
  Quality: 101.00      Length: 218
  Ratio: 0.971         Gaps: 11
  Percent Similarity: 47.706  Percent Identity: 22.477

alignment_block:
  US-09-525-867-1 x US-09-498-599-7  ..
  Align seg 1/1 to: US-09-498-599-7 from: 1 to: 3524
      8 GlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAl 24
      1046 GGTGTTCAGGCTGCGCCCATTCGGCAGCAACAGTCTCGCCGCGAT 1095
      24 aValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSers 41
      1096 TGTG..... 1099
      41 erThrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSer 57
      1100 .....CGTGCATTCCACCCCGCGGAA 1123
      58 Ser.ArgGlyGlyThrValValAlaLysLeuAspSerValAlaThrProAl 74
      1124 GTGAACAAACAGTCAAGTA.....TGGG 1149
      74 laArgArgSerLeuTrpProMetThrPheGlyLeuAlaCysCys... 89
      1150 CGGGGAATTCCTCGAGCTGGATTTTAAAGAGGAAGCTGCGACGCGGATG 1199
```

```
90 AlaVal..GluMetMetHisMetAlaAlaProArgTyrAspMetAspArg 105
   |||:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1200 GCTATGCCAAAGTGATGTCGGACGCGTTTCATCAAGCGGAATGGAA... 1246
106 PheGlyValValPheArgAlaSerProArgGlnSerAspValMetIleVa 122
   :..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1247 .....CTCTTTCGCCGCCAGCAAGAGGTCGATATCATGTGTAC 1287
122 lAlaGlyThrLeuThrAsnLysMetAlaProAlaLeu.....ArgLysV 137
   :..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1288 CACCGGCGTTATTCAGGCAACACCGCGCGGAAGCTAATTACCCGTGAA 1337
137 alTyrAspGlnMetProGluProArgTyrValValSerMetGlySerCys 153
   :..:..:..:..:..:..:..:..:..:..:~::~:~::~:~::~:~::~:~::~:
1338 TGGTTGACTCCATGAAGCGGCGAGTGTGATTGTGACCTG...GCAGCC 1384
154 AlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValValArgGlyCy 170
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1385 CAAACCGCGGC.....AACTG 1401
170 sAspArgIleValProValAspIleTyr..... 179
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1402 TGAATACACCGTCGCGGTGAATCTTCATACGGAAATGGTGTCAAAG 1451
180 .....IleProGlyCysProProThrAlaGluAlaLeu 190
1452 TGATTGGTTATACCGATCTCCGGGCGCTCTCGCGACGCAATCCTCACAG 1501
191 LeuTyrGly.....IleLeuGlnLeuGlnArgLysIleLysArgGluAr 205
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1502 CTTTAGCGCACAAACCTCGTTATCTGCTGAAACTGTTGTGCAAGAGAA 1551
205 g 205
1552 A 1552

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-178-252-26

seq_documentation_block:
; Sequence 26, Application US/09178252
; Patent No..6218188
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2
; CURRENT APPLICATION NUMBER: US/09/178,252
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/065,215
; EARLIER FILING DATE: 1997-11-12
; EARLIER APPLICATION NUMBER: 60/076,445
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 26
; LENGTH: 1965
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
; US-09-178-252-26

alignment_scores:
  Quality: 86.50      Length: 200
  Ratio: 0.911       Gaps: 8
  Percent Similarity: 47.500  Percent Identity: 26.500

alignment_block:
  US-09-525-867-1 x US-09-178-252-26/rev  ..
```

Align seg 1/1 to reverse of: US-09-178-252-26 from: 1 to: 1965

```

4 LeuSerAlaProGlyLeuArgGlyPheArg.....lleLeuGlyLe 17
|||||: ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1191 CTGGCGTAGGAGGCTCTCGTAGTGCTGCTGCTGCTGCTGCTGCT 1142
17 uArgSerValGlyProAlaVal.....GlnAlaArgGlyValHisGlnS 33
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1141 GCAGGGTGGCGGACAGTAGTCAGGAAGAGTGGTGGTGGCGGCG 1092
33 erValAlaThrAspGlyProSerThrGlnProAlaLeuProLysAla 49
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1091 GGTAGCGGTTCGGGACATGTTGAGCGGTTGAGCGCTCGAAGAGTG 1042
50 ArgAlaValAlaProLysProSerSerArgGlyGlyValVal..... 64
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1041 CGCGCGCTGTAGTAGGCGGTCTCCAGTCGAGAGGTGTTTACGGGT 992
65 .....AlaLysLeuAspAspLeuValAsnTrpAlaArgArgSers 78
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 TGTTCGCCAGCGCGGCGAGATGCCCTGGT.....GGCGGCGG.... 952
78 erLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMet 94
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
951 .....GTTGTA 946
95 HisMetAlaAlaProArgTyrAspMetAspArgPheGlyValValPheAr 111
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
945 CAC.....GATCGGTGGTGTAGATCTCGC 920
111 gAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeuThrA 128
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
919 CGGTGAGCTGGGGTTGGCGCGTTCGGGTACTGGCGCACGTTGTAGAAC 870
128 snLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProGluPro 144
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
869 GGAAGAGCGGCACAGTCATCGCATGAGGTCTCCCTCGCGGAA 820
145 ArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHi 161
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 GCGGTGGTGTAGTCCAGGACTC.....GGTGTGGTGGCGGCA 779
161 sTyrSerTyrSerValValArgGlyCysAspArgIleValProValAspI 178
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
778 CTTCTCCAGCGCGGTGTAGTCTCCAGCAGTGGTGGTGTAGAT. 730
178 leTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIle 194
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 .....GCGGTGGCTCTGCTGGCGGTT 706

```

seq_name: /cgn2_6/prodata/2/ina/6A_COMB.seq: US-09-036-987A-1

seq_documentation_block:
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:

```

; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dow AgroSciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA

```

```

ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 80161 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

```

alignment_scores:
Quality: 86.00 Length: 216
Ratio: 0.796 Gaps: 9
Percent Similarity: 50.000 Percent Identity: 22.685

alignment_block:

US-09-525-867-1 x US-09-036-987A-1 ..

```

Align seg 1/1 to: US-09-036-987A-1 from: 1 to: 80161
10 ArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAlaValG1 26
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6480 CGTGATTCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCG 6529
26 nAla...ArgGlyValHisGlnSerValAlaThrAspGlyProSerSert 42
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6530 TGCCACCGGTCTGCTCGTCGACGCGGCCGCCGCCGCCGCCGCCGCCG 6579
42 hrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSerSer 58
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6580 TCGATCCAGCCAGGTCGGG.....GATTCCTGTCGCGGTGTAATC 6602
59 ArgGlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaAr 75
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6603 .....GATTCCTGTCGCGGTGTAATC 6625
75 gArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysAlaValG 92
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6626 CGGTACGGATGCCAATATCCGTCCTGGCATCCGCCACCATCTCAGTAG 6675
92 luMetMetHisMetAlaProArgTyrAspMetAspArg..... 105
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6676 CGGTAGGCAAGTGTCTGCACAGAGTCACTCTGCGCGCCCTTTCCGCC 6725
106 PheGlyValValPheArg..AlaSerProArgGlnSerAspValMetIleV 122
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6726 AGCGTCTGGTTTCCGGCTCTGCATCCAGCGCAGCATCAGTC..... 6768
122 alAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgLysValTyr 138
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6769 .....TTCCGCCCTTGCCTTCAGGAGATGAGC 6795
139 AspGlnMetProGluProArgTyrValValSerMetGlySerCysAlaAs 155
|||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6796 GATGCCCGTGGCGAATCGCTTATGAGTCCCGGCGGACAGTGTGTGT 6845

```



```

; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-834-655-3

alignment_scores:
  Quality: 84.50      Length: 281
  Ratio: 0.741       Gaps: 16
  Percent Similarity: 40.569  Percent Identity: 20.996

alignment_block:
US-09-525-867-1 x US-08-834-655-3

Align seg 1/1 to: US-08-834-655-3 from: 1 to: 1488

25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
196 ATCGATGCGGGTTTGACCCAGCGTCATATCAGCACCTCGGCCCAACTC 245
41 rThGlnProHlaLeuProLysAlaArgAlaValAlaProLysProSer 57
246 GGCAAGCTGCTCGAGGCACTACCAAGCTCCCGAGTTCCACATCA 295
58 SerArgGlyGlu..... 61
296 AGGATCCGAGATGTCATCCCTGCCACTGCTTGGCGCTCCGGTCTC 345
62 TyrValValAlaLysLeuAspLeu.....Vala 72
346 CGTGTCTCTCGCAGTGGCATCGATCTGACTTGCGCTCGCTCTGTT 395
72 snTtpAlaArgArg.....SerSerLeu 79
396 CTGGCTGCGACCCAGATCGACAGTTTGAGATCCCTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysAlaValGluMetMethisMe 96
446 TGGCTGCGCTGTCTTACTGATCATGACGAGTA.....TT 480
96 talaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAla 113
481 GTCTGACCGGTGTCTGGGTCTGGCTCAGAGTGTGGTCACTGATCCTT 530
113 er.....ProArgGlnSerAspValMetileValAlaGlyThrLeuThr 127
531 CTGACCTCCAAAGACCTCAACACACAGTGTGT...GGATCTTGCAC 577
128 AsnLysMetAlaPro.....AlaLeuArgLys. 136
578 CGATGCTGTGGTCCCTACCTCCCTGGAGATCTCGCACTCGAAGCAC 627
137 .....ValTyrAspGlnMetProGlu..... 143
628 CACAGGCCACTGCCATATACCAAGACCAAGTCTTTGTGCCCAAGAC 677
144 .....ProArgTyrValValSerMetGlySer.....CysAlaAs 155
678 CCGCTCCCAAGTTGGCTGTGCTCCCAAGGAGAGAGTGTGCTGCGGTC 727
155 nGlyGlyGlyTyrThrHisTyrSerTyrSerValValArgGly..... 169
728 AGGAGGAGGACAT.....GTCCGTGACCTGGAT 756
170 .....CysAsp..... 171
757 GAGGAGGCTCCCATGTGTGACTTTTCTGGATGTTGATCCAGTCTTGT 806
172 .....ArgTyleValPr 175
807 CGGATGCGCCGCTGACTGATTATGAACGCTCTGCCAAGACTACGGCC 856

```

```

175 oValAspileTyrIlePro..... 181
857 GCTGACCTCGCACTTCCACACGTACTCGCCATCTTTGAGCCCCCAAC 906
182 .....GlycysProProThr 186
907 TTTTTCGACATTATTATCTCGGACCTCGGTGTGTGGCTGCCCTCGGTGC 956
187 AlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArg 199
957 CCTGATCTATGCTCCATCGAGTGTGCTCTTGGACCGT 995

seq name: /cqn2_6/ptodata/2/ina/6A_COMB.seq:US-08-834-033A-3

seq_documentation_block:
; Sequence 3, Application US/08834033A
; Patent No. 6075183
; GENERAL INFORMATION:
; APPLICANT: KNUDSON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,033A
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-300.USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-834-033A-3

```

```

alignment_scores:
  Quality: 84.50      Length: 281
  Ratio: 0.741       Gaps: 16
  Percent Similarity: 40.569  Percent Identity: 20.996

alignment_block:
US-09-525-867-1 x US-08-834-033A-3

Align seg 1/1 to: US-08-834-033A-3 from: 1 to: 1488

25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
:.....:
:.....:

```

```
196 ATCGATGCGGTTTGACCGAGGTGATATACAGCACTCGGCCCAAACTC 245
41 rThrGlnProAlaLeuProLys..AlaArgAlaValAlaProLysProSer 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246 GGCAAGCGCTGCTTGGAGCGCAACTACCAGCTCCCGGAGTTCAACATCA 295
58 SerArgGlyGlu..... 61
296 AGGAGATCGGAGAGTGCATCCCTGCCACTGCTTTGAGCGGTCCGGTCTC 345
62 .TyrValValAlaLysLeuAspLeu.....Vala 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 CGTGGTCTCTGCCACCTTGCCTGCATCTGACTTGGCGCTCGCTCTTGT 395
72 snTrpAlaArg.....SerSerLeu 79
396 CTTGGTTCGACCGAGATCGACAAGTTTGAGATCCCTTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHisMe 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
446 TGGCTGGCTGTTTACTGGATCATGACGGTA.....TT 480
96 tAlaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAlas 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 GTCTGACCGCTGCTGCTGGTGGTGCCTACAGAGTGTGGTCATCAGTCTT 530
113 er.....ProArgGlnSerAspValMetIleValAlaGlyThrLeuThr 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
531 CTGACCTCAAGACCTTCAACACACAGTTGGT....GGATCTTGCAC 577
128 AsnLysMetAlaPro.....AlaLeuArgLys. 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 CGATGCTCTTGGTCCCTACCACCTCTGAGAGATCTGCACTCGAAGCAC 627
137 .....ValTyrAspGlnMetProGlu..... 143
628 CACAAGGCCACTGGCCATATGACCAAGGACAGGTCTTTGTGCCCAAGAC 677
144 .....ProArgTyrValValSerMetGlySer.....CysAlaAs 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678 CCGCTCCAGGTGGTGTGCTCCCAAGGAGACGCTGCTGCTGCGCTTC 727
155 nGlyGlyTyrTyrHisTyrSerTyrSerValValArgGly..... 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
728 AGGAGGAGGACAT.....GTCCGTGCACCTGGAT 756
170 .....CysAsp..... 171
757 GAGGAGCTCCCATTTGACTTTGTTGATGGTGGATCCAGTTCTTCTT 806
172 .....ArgIleValPr 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
807 CGGATGCGCCGCGTACCTGATTATGAACCCCTGTGCCAAGACTACGGCC 856
175 oValAspIleTyrIlePro..... 181
857 GCTGGACCTCGCACTTCCACAGTACTGCCCATCTTTGAGCCCGGCAAC 906
182 .....GlyCysProProThr 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
907 TTTTTCGACATTATTATTCGGACCTCGGTGTGTTGGCTGCGCTCGGTGC 956
187 AlaGluAlaLeuTyrGlyIleLeuGlnLeuGlnArg 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
957 CCGTATCTATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-363-574-3

seq_documentation_block:

; Sequence 3, Application US/09363574
; Patent No. 6136574
; GENERAL INFORMATION:
; APPLICANT: KNUTZON, DEBORAH

```
; APPLICANT: MURKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,574
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-202 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-363-574-3
```

alignment_scores:
Quality: 84.50 Length: 281
Ratio: 0.741 Gaps: 16
Percent Similarity: 40.569 Percent Identity: 20.996

alignment_block:

US-09-525-867-1 x US-09-363-574-3 ..

Align seg 1/1 to: US-09-363-574-3 from: 1 to: 1488

```
25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 ATGATGCGGTTTGACCGAGGTGATATACAGCACTCGGCCCAAACTC 245
41 rThrGlnProAlaLeuProLys..AlaArgAlaValAlaProLysProSer 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246 GGCAAGCGCTGCTTGGAGCGCAACTACCAGCTCCCGGAGTTCAACATCA 295
58 SerArgGlyGlu..... 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 AGGAGATCGGAGAGTGCATCCCTGCCACTGCTTTGAGCGGTCCGGTCTC 345
62 .TyrValValAlaLysLeuAspLeu.....Vala 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 CGTGGTCTCTGCCACCTTGCCTGCATCTGACTTGGCGCTCGCTCTTGT 395
72 snTrpAlaArg.....SerSerLeu 79
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
396 CTTGGTTCGACCGAGATCGACAAGTTTGAGATCCCTTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHisMe 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


150 tGly.....SerCysAlaAsnG 156
:::|||||
1318 GTCTGCTGGGTACGTCGCCGCGAGACACGGTGCACGTGCTCGGA 1367
156 lYgLyGlyTyTyHisTyrSerValValArgGlyCysAspArg 172
|||||
1368 CCGCGGC.....GGTCCGGGAG 1387
173 lIeVal.....ProValAspIleTyrl 180
1388 TGCTGTCAACATCGTCTCCACGGGATCAAGTACAATCGGTCGGTGGT 1437
180 eProGlyCysProProThrAlaGluAlaLeuLeuTyrlGlylIleLeuGlnL 197
:::|||||
1438 GTGCTGACCCCAACAGATCAGGGGGTCTGCTCCGCGTCACACGAG 1487
197 euGlnArgLysIleLysArgGluArgArg 206
|||||
1488 GGCCCGGACTACCCAGCGAGCGCGACGA 1516

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-577-184-5

seq_documentation_block:
; Sequence 5, Application US/08577184
; Patent No. 5602014
; GENERAL INFORMATION:
; APPLICANT: MIZUMURA, YURIE
; APPLICANT: YU, FUJIO
; TITLE OF INVENTION: A REGULATORY FACTOR FOR EXPRESSION OF
; NITRILASE GENE AND A GENE THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P. O. NIKAKI, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/577,184
; FILING DATE: 22-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-337652
; FILING DATE: 28-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2941-004-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: RHODOCOCUS ERYTHROPOLIS
; STRAIN: SK92
; US-08-577-184-5

alignment_scores:

Quality: 82.50 Length: 243
Ratio: 0.809 Gaps: 9
Percent Similarity: 41.975 Percent Identity: 20.988
alignment_block:
US-09-525-867-1 x US-08-577-184-5
Align seg 1/1 to: US-08-577-184-5 from: 1 to: 2336
7 ProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValcylPr 23
|||:::|||||
1573 CCGAGGTGGTGGTATCGGTCCACCTGCCATCAGCGGTATGTGGACC 1622
23 oAlaValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProS 40
|:::|||||
1623 ACGAGCGGAGTTCATGGGATTCAGCGTCCATGCTCTGGACATCTCGA 1672
40 erSer.....ThrGln 43
|||:::
1673 GCGCTGGCGCGGACGCGCGCGGAGGACCACTCGAGCGCGAGGTG 1722
44 ProAlaLeuProLysAlaArgAlaValAlaProLysProSerSer..... 58
|||||:::|||||
1723 CCGCCATACCGCGGAGGAGCGGAGTCCGCGCAAGTCCACGTTCCTGCC 1772
59ArgGlyLeuTy-ValValAlaLysLeuAspAspLeuV 71
|||:::|||||
1773 GGCATGACGACGAGTCCAAACGCCCATGGCGTTATCGTCGGATTGAG 1822
71 alaTrpAlaArgSerSerLeuTrpProMetThrPheGlyLeuAla 87
|||:::|||||
1823 TGAGCTACTCGAGCGC.....TGG.....ACCTGGATGAAGAAC 1857
88 CysCysAlaValGluMetHisMetalAlaProArgTyrAspMetas 104
|||:::|||||
1858 GTGCTGAGTGGCGCTACCCCAAGTCCGCGAGCGCGAAGACACGTGATC 1907
104 pArgPheGlyValValPheArgAlaSerProArgGlnSerAspValMetI 121
|||:::|||||
1908 TCCTGTGTCGAGCGGTTCGATATAGCAAGATCGAAGCGCGGCTAT 1957
121 leValAlaGlyThrLeuThrAsn.....LysMetalAlaProAla 133
|||:::|||||
1958 CACTCTGCGAGGAGGAGATCGCTGCGAAGAGTTCGCCACCATCG 2007
134 LeuArgLysValTyraSpGlnMetProGluProArgTyrValValSerMe 150
|||:::|||||
2008 TGGAGA.....TGCTCGAGCCCATCGCCCGTGCACCGGTGAC 2048
150 tGly.....SerCysAlaAsnG 156
|||:::|||||
2049 GTCTGCTGCGGTACGTCGCCGCGAGACACCGGTGCGTGTGCTCGGA 2098
156 lYgLyGlyTyTyHisTyrSerTySerValValArgGlyCysAspArg 172
|||||
2099 CCGCGGC.....GGTCCGGGAG 2118
173 lIeVal.....ProValAspIleTyrl 180
1119 TGCTGTCAACATCGTCTCCACGGGATCAAGTACAATCGGTCGGTGGT 2168
180 eProGlyCysProProThrAlaGluAlaLeuLeuTyrlGlylIleLeuGlnL 197
|||:::|||||
2169 GTGCTGACCCCAACAGGATCAGGGGCTGCTCTCGCGGTGACACGAG 2218
197 euGlnArgLysIleLysArgGluArgArg 206
|||:::|||||
2219 GGCCCGGACTACCCAGCGAGCGCGACGA 2247

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-226-012-1

seq_documentation_block:

; Sequence 1, Application US/09226012

```

; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; TITLE OF INVENTION: SYNDROME GENE
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226,012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122,847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3477)
US-09-226-012-1

```

```

alignment_scores:
  Quality: 80.50      Length: 147
  Ratio: 1.134      Gaps: 6
  Percent Similarity: 48.299      Percent Identity: 27.211

alignment_block:
US-09-525-867-1 x US-09-226-012-1/rev ..

Align seg 1/1 to reverse of: US-09-226-012-1 from: 1 to: 3480

41 SerThrGlnProAlaLeuProLysAlaAlaValAlaProLysPro... 56
||| ||||| ||| ||||| ||||| ||||| ||||| |||||
2363 TCGATGGAGCCCGGAGATGAGTACAGCGCGGTGAGCAGCTCCCGCAGC 2314

57 .....SerSerArgGlyGluTyrValValAlaLysLeuAspA 69
||| ||| ||| |||||
2313 ATGCACCACTGTGTCCCTCGCGGTGATGTGTGTCTTG.....A 2273

69 sLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrPheGly 85
..... ||||| ||||| ||||| ||||| |||||
2272 ACTTCATGCCAGGCCCGAGCGACGCCCTTGTGGCCCTCGGAGCGGT 2223

86 LeuAlaCysCysAlaValGluMetMet.....HisMetAlaAlaPr 99
||| ||||| ||| ||| ||| ||||| ||||| |||||
2222 TTGCAGTGTGCAGCAGTGCAGCGGTTCAGGTGCAGCAGATGTCAGCCTG 2173

99 OArgTyrAspMetAspArgPheGlyValValPheArgAlaSerProArg 116
||| ||||| ||||| ||||| ||||| ||||| |||||
2172 CAGGCACCTCAGGAAGCCCTTCAGCAGCGGTTCATGTCGATGCCGTGG 2123

116 InSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaPro 132
||| ||| ||| ||||| ||||| ||||| |||||
2122 TGTAGGACCGAGC....GTGCTGGAAGTACTCTCAGGCGCGCTGCGCGAG 2077

133 AlaLeuArgLysValTyr.....AspGlnMetProGluProArgTyrVa 147
..... ||| ||||| ||||| ||||| ||||| |||||
2076 GGGATTGGGATCTGGTGGAGCGGATGAACCTCCCGCAGCCGACATCT 2027

147 lValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSert 164
;
2026 G..... 2026

164 yfSerValValArgGlyCysAspArgIleValProValAsp 177
||| ||||| ||||| ||||| ||||| |||||
2025 ..TGTTGGTACGGGCTGTGCCCGAGTACAGCCGCTGGAT 1987

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 19:53:05 ; Search time 15.62 Seconds
(without alignments)
306.863 Million cell updates/sec

Title: US-09-525-867-1
Perfect score: 1118
Sequence: 1 MAVLSAPGLRGFRILGRSS.....TLQLRKIKRRRLQIWR 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	718	64.2	182	3	US-09-261-471-5
2	88.5	7.9	510	3	US-08-801-344-8
3	88.5	7.9	510	4	US-09-498-599-8
4	70.5	6.3	319	4	US-08-836-075A-42
5	70.5	6.3	425	1	US-08-615-170-17
6	70.5	6.3	432	1	US-08-615-170-3
7	70.5	6.3	445	1	US-08-615-170-5
8	70.5	6.3	497	3	US-08-258-287B-37
9	70.5	6.3	497	3	US-08-368-704C-37
10	70	6.3	4302	3	US-08-558-136-5
11	69.5	6.2	2955	2	US-08-443-260-3
12	69.5	6.2	2955	3	US-08-442-805A-3
13	69.5	6.2	2955	3	US-08-443-900A-3
14	69.5	6.2	2955	4	US-08-444-818-124
15	69.5	6.2	2955	4	US-08-249-843-3
16	69.5	6.2	3011	2	US-08-833-678A-6
17	69.5	6.2	3011	4	US-08-444-818-177
18	69.5	6.2	3011	4	US-08-529-169A-6
19	69	6.2	457	4	US-09-124-238A-10
20	68	6.1	310	5	PCT-US95-04801-6
21	68	6.1	395	1	US-08-318-947A-18
22	68	6.1	395	2	US-08-795-303-18
23	68	6.1	748	2	US-08-997-080-154
24	68	6.1	748	2	US-08-997-362-154
25	68	6.1	748	4	US-08-995-855-154
26	68	6.1	888	4	US-09-268-140-4
27	68	6.1	993	4	US-09-060-410-4

28	67.5	6.0	2995	4	US-08-444-818-138	Sequence 138, App
29	67	6.0	3025	6	5223423-3	Patent No. 5223423
30	67	6.0	4303	2	US-08-460-751-2	Sequence 2, Appl
31	66.5	5.9	665	2	US-08-846-762-17	Sequence 17, Appl
32	66.5	5.9	665	2	US-08-846-762-94	Sequence 94, Appl
33	66.5	5.9	967	1	US-08-188-281B-13	Sequence 13, Appl
34	66.5	5.9	967	5	PCT-US94-07280-13	Sequence 13, Appl
35	66.5	5.9	967	5	PCT-US95-01087-13	Sequence 13, Appl
36	66.5	5.9	1648	1	US-08-188-281B-12	Sequence 12, Appl
37	66.5	5.9	1648	5	PCT-US94-07280-12	Sequence 12, Appl
38	66.5	5.9	1648	5	PCT-US95-01087-12	Sequence 12, Appl
39	66.5	5.9	3011	1	US-08-188-281B-1	Sequence 1, Appl
40	66.5	5.9	3011	1	US-08-453-552-1	Sequence 1, Appl
41	66.5	5.9	3011	2	US-08-710-637-1	Sequence 1, Appl
42	66.5	5.9	3011	3	US-08-811-566-20	Sequence 20, Appl
43	66.5	5.9	3011	4	US-09-014-416-1	Sequence 1, Appl
44	66.5	5.9	3011	5	PCT-US91-02225-10	Sequence 10, Appl
45	66.5	5.9	3011	5	PCT-US93-00907-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-261-471-5
; Sequence 5, Application US/09261471
; Patent No. 6025123
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/261,471
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/099,677
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0545 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: g2258365
US-09-261-471-5

Query Match 64.2%; Score 718; DB 3; Length 182;
Best Local Similarity 81.4%; Pred. No. 2.2e-76;

```

Db      291  --NCEYTVPGEIFTTENGVRKVIGYTDLPGRKLTQSSQLYGTNLVNLKLCKEK 342
RESULT          3
US-09-498-599-8
; Sequence 8 , Application US/09498599
; Patent No. 6303352
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF
; TITLE OF INVENTION: 1,2-PROPANEDIOL FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,599
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 8:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Pridine nucleotide transhydrogenase, subunit
; ORGANISM: A
; US-09-498-599-8

Query Match              7.9%; Score 88.5; DB 4; Length 510;
Best Local Similarity    23.7%; Pred. No. 0.079;
Matches 27; Conservative 23; Mismatches 35; Indels 29;

Qy   109 VFRASPRQSDVMIVAGITLTKMAPAL--RKVDMPDPRIYVMSGCANGGYYHYHSY
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   241 LFAAQAKEVDIIVTALIPGPKPAKPLITREMYDSMKAGSVIVDL-AAQNGG-----
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy   167 VRGCDRIVPVDIY-----IPGCPPTAEALLYG--ILQLORKTKRR 205
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db   291  --NCEYTVPGEIFTTENGVRKVIGYTDLPGRKLTQSSQLYGTNLVNLKLCKEK 342
RESULT          4
US-08-836-075A-42
; Sequence 42 , Application US/08836075A
; Patent No. 6180768
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: STUYVERS, LIEVEN

```

;; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
;; TITLE OF INVENTION: AGENTS AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
;; NUMBER OF SEQUENCES: 207
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ARNOLD, WHITE & DURKEE
;; STREET: P.O. BOX 4433
;; CITY: HOUSTON
;; STATE: TEXAS
;; COUNTRY: USA
;; ZIP: 77210-4433
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Microsoft Word 6.0 / ASCII text output
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/836,075A
;; FILING DATE: 21 Apr 1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/EP95/04155
;; FILING DATE: 23 Oct 1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: EP 94870166.9
;; FILING DATE: 21 Oct 1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: EP 93870076.7
;; FILING DATE: 28 Jun 1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KAMMERER, PATRICIA A.
;; REGISTRATION NUMBER: 29,775
;; REFERENCE/DOCKET NUMBER: INNS:004
;; INFORMATION FOR SEQ ID NO: 42:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 319 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-836-075A-42

Query Match 6.3%; Score 70.5; DB 4; Length 319;
Best Local Similarity 19.5%; Pred. No. 5.1;
Matches 45; Conservative 40; Mismatches 75; Indels 71; Gaps 14;
Qy 7 PGLRGFRILG---LRSSVGPVAVQARGVHQSVATDGPSSSTQPALPKARAVAPKPSRRGEYV 63
Db 25 PG--GGQIVGGVILLPRGRGLGVTRATKTSERSQPRGRQP-PKYR-----QPTGR---- 74
Qy 64 VAKLDDLNNARRSSLNPTMTGLACCAVEMHMAAPRYMDRFG-----VVVF 110
Db 75 -----SMQPGGYWPL-YGNEGCGWAGW-LLSPRGSRPNWGNDRRRSRNLGKVID 124
Qy 111 RASPRQSDVM---IVAGTL---TNKMAPALRKVD-----QMPD---PRVYVSMGSC 153
Db 125 TLTXLADLMGYIVLGGPLGGVAAALAHGVRAIEGVNATGNLPGCSFSLALALLSC 184
Qy 154 -----ANGGGYHY-----SYSWRGCDRIYVPIYIPGCPPTAE 189
Db 185 LTPATSAIQVKNASGIYHLTNDSCNNSIVEAETMI---LHLPGCVPCIK 232

RESULT 5
US-08-615-170-17
; Sequence 17, Application US/08615170
; Patent No. 5776776
; GENERAL INFORMATION:
; APPLICANT: ORDAHL, Charles P.
; APPLICANT: AZAKIE, Anthony
; APPLICANT: MAR, Janet H.
; APPLICANT: FARRANCE, Iain K.G.
; APPLICANT: HALL, Deborah E.
; APPLICANT: STEWART, Alexandre F.R.

APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-615-170-17

Query Match 6.3%; Score 70.5; DB 1; Length 425;
Best Local Similarity 23.3%; Pred. No. 7.8;
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;
Qy 3 VLSAPGLRGFRIL-GLRSSVGPVAVQ-----RGVHQSV-----TDGPSSTQPALPKAR 50
Db 130 IISATAPHSKMLPGLPRSAYPVSGFWQGLPGQAGSSQDVKPFQQPYALQPSLPLPG 189
Qy 51 AVAP-----KPSSRGVYVAKLDDLNNWARRSSLPMTGLACCAVEMHMAAPRYD 102
Db 190 FDSPTGLPPSSSPFAGGRRVAS-----SKLWLEFS-----AFLEQQQODDTN 234
Qy 103 MDREGVYFRASPRQSDVMIVAGTITNKNMAPALRKVYDQMEPR 145
Db 235 KHLFVHIGQSNPSYSDPYL-----EAVDIRQYDKFPEKK 269

RESULT 6
US-08-615-170-3
; Sequence 3, Application US/08615170
; Patent No. 5776776
; GENERAL INFORMATION:
; APPLICANT: ORDAHL, Charles P.
; APPLICANT: AZAKIE, Anthony
; APPLICANT: MAR, Janet H.
; APPLICANT: FARRANCE, Iain K.G.
; APPLICANT: HALL, Deborah E.
; APPLICANT: STEWART, Alexandre F.R.
; APPLICANT: LARKIN, Sarah B.

;; TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
;; NUMBER OF SEQUENCES: 32
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend Kourie and Crew
;; STREET: Stewart Street Tower, One Market Plaza
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: US
;; ZIP: 94105-1493
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/615,170
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/01526
;; FILING DATE: 06-FEB-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/191,493
;; FILING DATE: 04-FEB-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Heslin, James M.
;; REGISTRATION NUMBER: 29,541
;; REFERENCE/DOCKET NUMBER: 2307U-053120
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 432 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-615-170-3

Query Match 6.3%; Score 70.5; DB 1; Length 432;
Best Local Similarity 23.3%; Pred. No. 8;
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;

QY 3 VLSAPGLRGRFRIL-GLRSSYGVPAVQA-----RGVHQSVA--TDGPSSTQPALPKAR 50
Db 137 IISATAFHSKMAPLGLPRSAYPVSGFWOGALPGQAGSSQDVKPTQOQYALQPSLPLG 196
QY 51 AVAP-----KPSRGEYVYAKLDLVNWARSSLWMTFGLACCAVEMMHMAAPRYD 102
Db 197 FDSPTGLPPSSSTPAWQGRVAS-----SKLWLEFS----AFLEQQQDDQDTYN 241
QY 103 MDRFGVVFRAPSQSDVMIVAGTLTNKMAPALRKVYDQMPEPR 145
Db 242 KHLFVHIGQSNPSYSDPYL-----EAVDIRQIYDKPEKK 276

RESULT 7
US-08-615-170-5
; Sequence 5, Application US/08615170
; Patent No. 5776776
; GENERAL INFORMATION:
; APPLICANT: ORDAHL, Charles P.
; APPLICANT: AZAKIE, Anthony
; APPLICANT: MAR, Janet H.
; APPLICANT: FARRANCE, Iain K.G.
; APPLICANT: HALL, Deborah E.
; APPLICANT: STEWART, Alexandre F.R.
; APPLICANT: LARKIN, Sarah B.
; TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
; NUMBER OF SEQUENCES: 32

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend Kourie and Crew
;; STREET: Stewart Street Tower, One Market Plaza
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: US
;; ZIP: 94105-1493
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/615,170
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/01526
;; FILING DATE: 06-FEB-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/191,493
;; FILING DATE: 04-FEB-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Heslin, James M.
;; REGISTRATION NUMBER: 29,541
;; REFERENCE/DOCKET NUMBER: 2307U-053120
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 445 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-615-170-5

Query Match 6.3%; Score 70.5; DB 1; Length 445;
Best Local Similarity 23.3%; Pred. No. 8.3;
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;

QY 3 VLSAPGLRGRFRIL-GLRSSYGVPAVQA-----RGVHQSVA--TDGPSSTQPALPKAR 50
Db 150 IISATAFHSKMAPLGLPRSAYPVSGFWOGALPGQAGSSQDVKPTQOQYALQPSLPLG 209
QY 51 AVAP-----KPSRGEYVYAKLDLVNWARSSLWMTFGLACCAVEMMHMAAPRYD 102
Db 210 FDSPTGLPPSSSTPAWQGRVAS-----SKLWLEFS----AFLEQQQDDQDTYN 254
QY 103 MDRFGVVFRAPSQSDVMIVAGTLTNKMAPALRKVYDQMPEPR 145
Db 255 KHLFVHIGQSNPSYSDPYL-----EAVDIRQIYDKPEKK 289

RESULT 8
US-08-258-287B-37
; Sequence 37, Application US/08258287B
; Patent No. 6083735
; GENERAL INFORMATION:
; APPLICANT: Yuan, Junying
; APPLICANT: Miura, Masayuki
; TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005


```
;
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-658-136-5

Query Match      6.3%; Score 70; DB 3; Length 4302;
Best Local Similarity 23.3%; Pred. No. 2.8e+02;
Matches 58; Conservative 24; Mismatches 57; Indels 110; Gaps 16:

QY 3 VLSAPGLRGRILGRSSVGAQARGVHQSVDGPGSTQPALPKARAVAPKPSRGEY 62
DB 707 VMLPG-----DLVGLQHDAGPGL---LHCSFPGHPG-----PRAPVLSANASSWLPH 753
QY 53 VYAKDDLVNWRSSLWPMFTGLACCAVEVMHMAAP-----RYDM 103
DB 754 LPAQLE-----GTW-----ACPACALRLIAATEQLTVLLGLRPNPGLRLPGRYEV 798
QY 104 D-----RGVVFRAASPROSDVMIVAGILTNTKMAPALRKVYDQMPER- 145
DB 799 RAENVGVSRNLSGSDVV---SP-----LVYIY---PAPRDGRL 836
QY 146 YVVMGSC-----ANG-----GGYHYHSYVVRGCDRVP--VDIYIPGCPPTAE 188
DB 837 YVPTNGSALVQVDSGANATATARWPG-----SVSARENVCPALVATFVPGCPWEIN 890
QY 189 ALLYGILOL 197
DB 891 DTLSVVAL 899

RESULT 11
US-08-443-260-3
; Sequence 3, Application US/08443260
; Patent No. 5942234
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDUM, KENT B.
; APPLICANT: GERVASE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ANTIGENIC PROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,260
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154,006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-658-136-5

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Arg or Lys"
OTHER INFORMATION: this location; Xaa = Arg or Lys"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Asn or Thr"
OTHER INFORMATION: this location; Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Thr"
OTHER INFORMATION: this location; Xaa = Ile or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 334
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Met or Val"
OTHER INFORMATION: this location; Xaa = Met or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Leu"
OTHER INFORMATION: this location; Xaa = Ile or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Asn or Tyr"
OTHER INFORMATION: this location; Xaa = Asn or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Pro or Ser"
OTHER INFORMATION: this location; Xaa = Pro or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Leu or Pro"
OTHER INFORMATION: this location; Xaa = Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Cys or Tyr"
OTHER INFORMATION: this location; Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1471
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Glu or Gly"
OTHER INFORMATION: this location; Xaa = Glu or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = His or Leu"
OTHER INFORMATION: this location; Xaa = His or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1949
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = His or Leu"
OTHER INFORMATION: this location; Xaa = His or Leu"
```



```

; OTHER INFORMATION: this location; Xaa = Leu or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1454
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1471
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1677
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1948
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = His or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1949
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2021
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2385
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2386
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ala or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2502
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2590
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2921
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Gly"
;
; US-08-442-805A-3

```

```

Query Match      6.2%; Score 59.5; DB 3; Length 2955;
Best Local Similarity 20.8%; Pred. No. 1.8e+02;
Matches 48; Conservative 29; Mismatches 75; Indels 79; Gaps 14;

QY 7 PGLGRFRILG--LRSSVGPAVQARGVSHVATDGPSTQPALPKARAVAPKPSRGEYV 63
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 PG--GGGIVGGVLLPRRGRLGVRATKTSRSQPRGRQPIPKAR-----RPGR---- 74
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 64 VAKLDLVNWARSSLPWMTFGLACCAVEMHMAAPRYDMDFGVVFRASPRQ----- 116

```

```

Db 75 -----TWAQPGYBPL-YGNEGCGWAGW-LLSPRGSRPSMG-----PTDPRRRNRNLGK 121
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 117 -----SDVM-----IVAGTLTNKMAPALRKRYD-----QMPE---PRYVWSM 150
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 VIDLTGCFADLMGYIPLVGAPLGSAARALAHGVRYLEDGVYATGNLPGCSFSXFLAL 181
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 151 GSC-----ANGGGYHY-----SYSVVRGCDRIVPVVDIYIPGCPP 185
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 LSCLTVPASAYQVRNSTGLYHTNDPCNPSSIVYEADAI-----LHTPGCV 228
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-08-443-900A-3
; Sequence 3, Application US/08443900A
; Patent No. 6074852
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDUM, KENT B.
; APPLICANT: GERVAISE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-May-1995
; CLASSIFICATION: 435
; APPLICATION NUMBER: US/08/443.900A
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Lys"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at

```

```

; LOCATION: 2502
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2690
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2921
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Gly"
US-08-0443-900A-3

```

```

Query Match          6.2%; Score 69.5; DB 3; Length 2955;
Best Local Similarity 20.8%; Pred. No. 1.8e+02;
Matches 48; Conservative 29; Mismatches 75; Indels 79; Gaps

      QY      7  PGLRGERILG--LRRSVGPAVGARGVHGVSATGDCPSSTOPALPKARAVAPKPPSSRGEYV 63
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      25  PG--GGQIVGVGVI LLPRGPRGLGVRAITRTSRSQRRRRQPIKAR-----RPEGR---- 74

      QY      64  VAKLDDL VNNARRSLMPTTGLACCAVEMHMAAPRYDMDRFGVVFRASPRQ-----116
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      75  -----TWAQPGYPWPL-YGNEGCGGAGW--LLSPGRSRPSWG---PTDPRRRSRNLGK 121

      QY      117 -----SDVM-----IVAGTLTNKMAPALKRYD-----OMPE--PRYVWSM 150
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      122  VIDTLTCGFADLMGYIPLVGAPGLSRAARALAHGVRLGEGVNYATGNLPGCSXFLLAL 181

      QY      151  GSC-----ANGGGYYHY-----SYSVVRGCDRIVFDIYIPGCCPP 185
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      182  LSCITLTPASAYQVNRSTGLYHTVNDPCNSSIIVYEADA1-----LHTPGCVF 228

```

RESULT 14
US-08-444-818-124
Sequence 124, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 124:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Duplication
LOCATION: 9
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Lys or Arg"
FEATURE:
NAME/KEY: Duplication
LOCATION: 11
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Asn or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 176
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Ile or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 334
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Met or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 603
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Leu or Ile"
FEATURE:
NAME/KEY: Duplication
LOCATION: 848
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Tyr or Asn"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1276
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Pro or Leu"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1454
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Cys or Tyr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1471
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1977
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Glu or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1948
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Leu or His"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1949
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Ser or Cys"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2021
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Gly or Val"
FEATURE:
NAME/KEY: Duplication
```

```
LOCATION: 2349
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2385
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Tyr or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2386
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Ser or Ala"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2502
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Leu or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2690
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2921
OTHER INFORMATION: /note= "A heterogeneity exists at
OTHER INFORMATION: Xaa which is either Arg or Gly"
US-08-444-818-124
```

Query Match 6.2%; Score 69.5; DB 4; Length 2955;
Best Local Similarity 20.8%; Pred. No. 1.8e+02;
Matches 48; Conservative 29; Mismatches 75; Indels 79; Gaps 14;

```
QY 7 PGLRGFRILG---LRSSVGPVQVARGVHQSVDGPGSSTQPALPKARAVAKPSSRGYV 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 PG-GGQIVGGVLLPFRGLGVRATRTSRSQPRGRQPIPKAR----RPEGR---- 74

QY 64 VAKLDDLNVNARRSSLWPMTFGLACCAVEMMHMAAPRYDMDRFGVVFRASPRQ----- 116
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 75 -----TWAQPGYEWPL-YGNEGCGWAGW-LLSPRGRSPSWG---PTDPRRSRNLGK 121

QY 117 -----SDVM-----IVAGLTNKNMAPALRKVD-----QMPE---PRYVVSM 150
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 VIDTLTCGFADLMGYIPLVGSAPLGSAAARALAHGVRLVDGVNVAATGNLPGGSFXLLAL 181

QY 151 GSC-----ANGGGYHYH-----SYSVVRCGRIVPDVIYIPGCPP 185
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 LSCLTVPASAYQVRNSTGLYHVTNDCPNSSIVYEADAI----LHTPGCV 228
```

```
RESULT 15
US-08-249-843-3
; Sequence 3, Application US/08249843
; Patent No. 6274148
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDIUM, KENT B.
; APPLICANT: GERVASE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,843
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0154,003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 665-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Arg or Lys"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 334
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Met or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Asn or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Pro or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site

LOCATION: 1471
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Glu or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = His or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1949
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Cys or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2021
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Gly or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2349
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2385
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Phe or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Phe or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Gly or Arg"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2921
OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Arg or Gly"
US-08-249-843-3
Query Match 6.2%; Score 69.5; DB 4; Length 2955;
Best Local Similarity 20.8%; Pred. No. 1.8e+02;
Matches 48; Conservative 29; Mismatches 75; Indels 79; Gaps 14;
QY 7 PGLRGRILG---LRSSVGPVQARGVHQSVATDGPSSSTQPALPKARAVAPKPSRGEYV 63
Db 25 PG--GGQIVGVYLLPRRGRIGRVATRTKTSERSQPRRRQPIPRK---RPEGR----- 74
QY 64 VAKLDDLNVNARRSSLWPMTEFGLACCAVEMHMAAPRYDMDFGVVFRASPRQ----- 116
Db 75 -----TWAQGPVWPL-YGNECGWAGW-LLSPGRSRPSWG---PTDPRRSRLGK 121
QY 117 -----SDVM-----IVAGTLTKMAPALARKYVD-----QMPE---PRYVSM 150
Db 122 VIDLTCGFADLMGYIPLVGAPLGSAAARALAHGVRLVEDGVNYATGNLPGCSFXFLAL 181
QY 151 GSC-----ANGGGYHY-----SYSVRGCDRIVPVDIVIPGCPP 185

Db 182 USCLTPASAYQVRNSTGLYHVTNDCPNSSIVYEADAI-----LHTPGCVP 228

Search completed: February 12, 2002, 20:40:47
Job time: 2862 sec

alignment_scores:
 Quality: 1118.00 Length: 213
 Ratio: 5.249 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-525-867-1 x AAI60566 ..

Align seg 1/1 to: AAI60566 from: 1 to: 752

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
 |||||
 5 ATGGCGGTGCTGTCAGCTCTGCGCTCGCGGGCTTCGGATCTTGGTCT 54
 |||||
 17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
 |||||
 55 CGCGTCCAGCGTGGCGCGGCTGTGCAGGCACGAGGTGTCCATCAGAGCG 104
 |||||
 34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
 |||||
 105 TGGCCACGATGGCCCAAGCAGCACCAGCTGCCCTGCCAAGGCCGAGA 154
 |||||
 51 AlaValAlaProLysProSerSerArgGlyGlyValValAlaLysLe 67
 |||||
 155 GCCGTGGCTCCCAACCCAGCAGCGCGGGGAGTATGTGGTGGCCAAAGCT 204
 |||||
 67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
 |||||
 205 GGATGACCTCTCAACTGGGCGCGCGGAGTCTCTGTGGCCCATGACCT 254
 |||||
 84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
 |||||
 255 TCGGCTGGCTGTGCGCGGTGGAGATGACATGCGCAGCAGCCCGCG 304
 |||||
 101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnse 117
 |||||
 305 TACGACATGGACCGCTTTGGCGTGGTCTCCGCGCGAGCGCGCGCAGTC 354
 |||||
 117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
 |||||
 355 CGAGCTCATGATCGTGGCGCGGCACACTCACCAACAGATGGCCCGCAG 404
 |||||
 134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
 |||||
 405 TTCGCAAGGTCTACGACACAGATGCGGAGCGCGGCTACGTGTCCATG 454
 |||||
 151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVal 167
 |||||
 455 GGGAGCTGGCGCAACGAGGAGGAGGCTACTACCATATTCCTACTCGGTGT 504
 |||||
 167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
 |||||
 505 GAGGGGCTGCAGCGCATCGTGGCCGCTGGACATCTACATCCCGAGGTGCC 554
 |||||
 184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuArgLys 200
 |||||
 555 CACCTACGGCGGAGGCGCTCTCTACGGCATCTCGAGTGTGAGTGGAGGAG 604
 |||||
 201 lIleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
 |||||
 605 ATCAAGCGGGGAGCGGAGGCTGCAGATCTGTGTACCGCAGG 643

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58780

seq_documentation_block:

ID AAI58780 standard; cDNA; 772 BP.

XX

AC AAI58780;

XX

DT 22-OCT-2001 (first entry)

XX

Human polynucleotide SEQ ID NO 983.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI: 2001-442253/47.

P-PSDB: AAM39624.

Novel nucleic acids and polypeptides, useful for treating disorders
 such as central nervous system injuries -

Claim 1; SEQ ID NO 983; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369), and
 the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 immunosuppressant and cytostatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 localised neuropathies and central nervous system diseases, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 utilisation of the activities such as: Immune system suppression,
 Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 assays for receptor activity, arthritis and inflammation, leukaemias and
 C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed
 specification.

Sequence 772 BP; 136 A; 265 C; 245 G; 126 T; 0 other;

alignment_scores:

Quality: 1118.00 Length: 213
 Ratio: 5.249 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-525-867-1 x AAI58780 ..

Align seg 1/1 to: AAI58780 from: 1 to: 772

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
 |||||
 24 ATGGCGGTGCTGTCAGCTCTGCGCTCGCGGGTTCGGATCTTGGTCT 73

```

17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnServ 34
18 |
19 |
20 |
21 |
22 |
23 |
24 |
25 |
26 |
27 |
28 |
29 |
30 |
31 |
32 |
33 |
34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
35 |
36 |
37 |
38 |
39 |
40 |
41 |
42 |
43 |
44 |
45 |
46 |
47 |
48 |
49 |
50 |
51 AlaValAlaProLysProSerSerArgGlyGlyValValAlaLysLe 67
52 |
53 |
54 |
55 |
56 |
57 |
58 |
59 |
60 |
61 |
62 |
63 |
64 |
65 |
66 |
67 |
68 |
69 |
70 |
71 |
72 |
73 |
74 GCGCTCCAGCGTGGCCCGCGCTGTGCAGCGACAGAGTGTCCATCAGAGCG 123
75 |
76 |
77 |
78 |
79 |
80 |
81 |
82 |
83 |
84 |
85 |
86 |
87 |
88 |
89 |
90 |
91 |
92 |
93 |
94 |
95 |
96 |
97 |
98 |
99 |
100 |
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
102 |
103 |
104 |
105 |
106 |
107 |
108 |
109 |
110 |
111 |
112 |
113 |
114 |
115 |
116 |
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
118 |
119 |
120 |
121 |
122 |
123 |
124 |
125 |
126 |
127 |
128 |
129 |
130 |
131 |
132 |
133 |
134 CGAGCTCATGTCGTGGCGCGGCACACTCACCACCAAGATGGCCCGCAGCGC 423
135 |
136 |
137 |
138 |
139 |
140 |
141 |
142 |
143 |
144 |
145 |
146 |
147 |
148 |
149 |
150 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
151 |
152 |
153 |
154 |
155 |
156 |
157 |
158 |
159 |
160 |
161 |
162 |
163 |
164 |
165 |
166 |
167 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValva 167
168 |
169 |
170 |
171 |
172 |
173 |
174 |
175 |
176 |
177 |
178 |
179 |
180 |
181 |
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 IleLysArgGluArgArgLeuGlnIleTyrTyrArgArg 213
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 ATCAACGGGGAGCGGAGGTGCAGATCTGGTACCGCAGG 662
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 |
782 |
783 |
784 |
785 |
786 |
787 |
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 |
842 |
843 |
844 |
845 |
846 |
847 |
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 |
902 |
903 |
904 |
905 |
906 |
907 |
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 |
962 |
963 |
964 |
965 |
966 |
967 |
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |

```

seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC59553

seq_documentation_block:

ID AAC59553 standard; CDNA; 785 BP.

XX AC AAC59553;

XX DT 15-FEB-2001 (first entry)

XX DE Human secreted protein cDNA sequence #47.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

XX KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

XX KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX KW neurological disease; infection; human; secreted protein; ss.

XX OS Homo sapiens.

XX PN WO200055352-A2.

XX PD 21-SEP-2000.

XX XX 09-MAR-2000; 2000WO-US06044.

XX 12-MAR-1999; 99US-0124099.
 PR 03-DEC-1999; 99US-0168661.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-602124/57.
 DR P-PSDB; AAB34261.
 XX Novel human secreted proteins useful for diagnosis, prevention and
 PT treatment of disorders including neurological, cell proliferative,
 PT cardiovascular, autoimmune and inflammatory disorders and microbial
 PT infections.
 XX Claim 1: Page 345; 383pp; English.
 XX The invention relates to the isolation of genes AAC59507-c59556 encoding
 CC 50 human secreted proteins AAB34218-B34264. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (AAC59498) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 XX Sequence 785 BP; 173 A; 255 C; 236 G; 121 T; 0 other;

alignment_scores:

Quality: 1065.00 Length: 209
 Ratio: 5.120 Gaps: 0
 Percent Similarity: 99.522 Percent Identity: 98.565

alignment_block:

US-09-525-867-1 x AAC59553

Align seg 1/1 to: AAC59553 from: 1 to: 785

5 SerAlaProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerVa 21
 18 ACAGTCTCTGGCTGCGCGGCTTCCGGATCTCTGGTCTGGCTCCAGCGT 67
 21 lGlyProAlaValGlnAlaArgGlyValHisGlnSerValAlaThrAspG 38
 68 GGGCCCGGCTGTGCAGGACGAGGTCTCCATCAGAGCGTGGCCACCGATG 117
 38 lYProSerSerThrGlnProAlaLeuProLysAlaArgAlaValAlaPro 54
 118 GCCCAAGCAGCACCAGCTGCCCTGCCAAAGGCCAGAGCGTGGCTCCC 167
 55 LysProSerSerArgGlyGlyTyrValValAlaLysLeuAspAspLeuVa 71
 168 AAACCCAGCAGCCGGCGAGTATGTGTGGTGGCCAGCTGGATGACCTCGT 217
 71 lAsnTrpAlaArgArgSerSerLeuTrpProMetThrPheGlyLeuAlaC 88
 218 CAACGGGCGCGCGAGTTCTCTGTGGCCCATGACCTTCCGCTGGCCT 267
 88 YScysAlaValGluMetHisMetAlaAlaProArgTyrAspMetAsp 104

PR	06-MAY-1999;	9905-01324887
PR	07-MAY-1999;	9905-01328863
PR	11-MAY-1999;	9905-01342563
PR	14-MAY-1999;	9905-01342119
PR	14-MAY-1999;	9905-01342118
PR	14-MAY-1999;	9905-01342219
PR	14-MAY-1999;	9905-01342211
PR	14-MAY-1999;	9905-01342170
PR	18-MAY-1999;	9905-01347668
PR	20-MAY-1999;	9905-01349411
PR	20-MAY-1999;	9905-01351124
PR	21-MAY-1999;	9905-01353553
PR	24-MAY-1999;	9905-01356629
PR	25-MAY-1999;	9905-01360211
PR	25-MAY-1999;	9905-01363992
PR	28-MAY-1999;	9905-01367882
PR	01-JUN-1999;	9905-01372222
PR	03-JUN-1999;	9905-01375228
PR	04-JUN-1999;	9905-01377502
PR	07-JUN-1999;	9905-01377702
PR	08-JUN-1999;	9905-01380909
PR	10-JUN-1999;	9905-01385401
PR	10-JUN-1999;	9905-01388847
PR	14-JUN-1999;	9905-01391119
PR	16-JUN-1999;	9905-01394552
PR	17-JUN-1999;	9905-01394553
PR	17-JUN-1999;	9905-01394922
PR	18-JUN-1999;	9905-01394554
PR	18-JUN-1999;	9905-01394555
PR	18-JUN-1999;	9905-01394556
PR	18-JUN-1999;	9905-01394557
PR	18-JUN-1999;	9905-01394558
PR	18-JUN-1999;	9905-01394559
PR	18-JUN-1999;	9905-01394650
PR	18-JUN-1999;	9905-01394661
PR	18-JUN-1999;	9905-01394662
PR	18-JUN-1999;	9905-01394663
PR	18-JUN-1999;	9905-01397500
PR	18-JUN-1999;	9905-01397501
PR	21-JUN-1999;	9905-01398177
PR	22-JUN-1999;	9905-01398999
PR	23-JUN-1999;	9905-01403553
PR	23-JUN-1999;	9905-01403554
PR	24-JUN-1999;	9905-01403554
PR	24-JUN-1999;	9905-01406955
PR	28-JUN-1999;	9905-01408923
PR	29-JUN-1999;	9905-01409991
PR	30-JUN-1999;	9905-01412887
PR	01-JUL-1999;	9905-01418482
PR	01-JUL-1999;	9905-01421554
PR	02-JUL-1999;	9905-01420555
PR	06-JUL-1999;	9905-01423990
PR	08-JUL-1999;	9905-01428003
PR	09-JUL-1999;	9905-01429200
PR	12-JUL-1999;	9905-01432977
PR	13-JUL-1999;	9905-01435452
PR	14-JUL-1999;	9905-01436204
PR	15-JUL-1999;	9905-01440055
PR	16-JUL-1999;	9905-01440085
PR	16-JUL-1999;	9905-01440086
PR	19-JUL-1999;	9905-01443251
PR	19-JUL-1999;	9905-01443331
PR	19-JUL-1999;	9905-01443332
PR	19-JUL-1999;	9905-01443333
PR	19-JUL-1999;	9905-01443334
PR	19-JUL-1999;	9905-01443335
PR	20-JUL-1999;	9905-01443532
PR	20-JUL-1999;	9905-01443532
PR	20-JUL-1999;	9905-01448884
PR	21-JUL-1999;	9905-01448814
PR	21-JUL-1999;	9905-01450886
PR	21-JUL-1999;	9905-01450885
PR	22-JUL-1999;	9905-01450887
PR	22-JUL-1999;	9905-01450889

```
PR 22-JUL-1999; 99US-01451192.
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-01452118.
PR 23-JUL-1999; 99US-01452181.
PR 26-JUL-1999; 99US-01452224.
PR 26-JUL-1999; 99US-01452276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-01459118.
PR 27-JUL-1999; 99US-01459119.
PR 28-JUL-1999; 99US-0145911.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147433.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-01493175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159594.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160757.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 737.50      Length: 214
    Ratio: 4.120        Gaps: 3
Percent Similarity: 83.645 Percent Identity: 66.822

alignment_block:
US-09-525-867-1 x AAC43639 ..

Align seg 1/1 to: AAC43639 from: 1 to: 911

7 ProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyPr 23
   ||| ||||| :||| :||| :||| :||| :||| :||| :||| :|||
65 CCGTCCACGAGGTGAGGAACCCCTAGCCGACCATGGCTGCTGCC 114
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
23 oAlaValAlaAlaArgGly.Val.....HisGlnSer 33
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
115 CGCGCACCGCGGCTGGCCCTCTTCACCCCGCGGCGTACTCCGCT 164
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
34 ValAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaAr 50
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
165 GCGGTGCTGCGGCGGCGCTCCCGAGCTCCAGCG...CCATACGGGG 211
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
50 gAlaValAlaProLysProSerSerArgGlyGlnTyrValValAlaLysL 67
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
212 CGCCCCCCCCCGGCGGTGCGAAGACGCGCGAGTCTGTCTCTAAGG 261
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
67 euAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThr 83
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
262 TGGACGACCTGATGAACCTGGGCGGTAAAGGCTCATCTGCGCCATGACC 311
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
84 PheGlyLeuAlaCysCysAlaValGlnMetHisMetAlaAlaProAr 100
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
312 TTCGGCTCGCTGCTGCGGTGGATGATGACACGCGGCGGTCCCG 361
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
100 gTyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnS 117
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
362 CTACGACTTCGACCGGTTCCGGGTCATCTCCGTCGCGCGGCCAGT 411
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
117 exAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAla 133
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
412 CCGATTGCTATGCTCGCGCGGCGCTCACCACACAGATGGCTCCAGCC 461
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
134 LeuArgLysValTyrAspGlnMetProGluProArgTyrValValSerMe 150
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
462 CTCGCCAAGGTTTATGACCAATGCTGAGCCTAGTGGGTATTTCAT 511
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
150 tGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValV 167
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
512 GGGCAGCTGTGCCAACGGTGGTGGATACTACCAATTACTCTCTGTTG 561
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
167 aArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCys 183
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
562 TACGTGGATGTACCGTATAGTCCCTGTGGACATCTAGTCCCTGGGTCC 611
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
184 ProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLy 200
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
612 CCACCAACTGCTGAGGACTGCTGTACGGCGCTTCTTACGCTCCAAAGAA 661
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```


PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140698.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156456.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 725.00      Length: 198
    Ratio: 4.290        Gaps: 2
    Percent Similarity: 85.354    Percent Identity: 68.182

alignment_block:
US-09-525-867-1 x AAC38037 ..

Align seg 1/1 to: AAC38037 from: 1 to: 976

24 AlaValGlnAlaArgGlyVal.....HisGlnSerValalathras 37
112 GCGGTCGCGCGTGGCTGCTCTCACCTCCACACATCTCCCGCTCT 161
37 PglyProSer.....ThrGlnProAlaLeuProLysa 49
162 CTCCTCCCTCCACATCAGGACTTCTTACACACAGCTGGTCTCTCGA 211
49 laArgAlaValAlaProLysProSerArgGlyGluTyrValValAla 65
212 COTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261
66 LysLeuAspLeuValAlaAsnTTPAlaArgSerSerLeuTrpProme 82
262 AAGGTCGATGATCATCAAGTGGCTGGTACCGGATCGATCTGGGCTAT 311
82 tThrPheGlyLeuAlaCysAlaValAlaGluMetHisMetAlaAlaP 99
312 GACCTTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
99 roArgTyrAspMetAspArgPheGlyValValPheArgAlaSerProArg 115
362 CFCGTCGATGATCTGATGATGATGATGATGATGATGATGATGATGAT 411
116 GlnSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaP 132
412 CAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 461
132 oAlaLeuArgLysValTyrAspGlnMetProGluProArgTyrValVal 149
462 GCGCTCTCGCAAGGTTTATGACCATGCTGACCAAGGTTGATGATTT 511

```

```

149 erMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSer 165
512 CAATGGGAAGTTGTGCCAATGGAGGTGGATACATCTACTACTCTACTCG 561
166 ValValArgGlyCysAspArgIleValProValAspIleTyrIleProG 182
562 GTGGTTGAGGATGTGACAGAAATGTCCAGTCGACATATACGTCCCGGG 611
182 yCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnA 199
612 GTGCCACCAACCGCTGAGGCTTTGCTCTATGGACTACTCCAGCTTCAGA 661
199 rGlyIleLysArgGluArgLeuGlnIleTrpTyrArgArg 213
662 AGAAATCAACAGCGCAAGGATTCTTGATTGCTGGTGGAAACAAG 705

seq_name: ./SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC41785
seq_documentation_block:
ID_ AAC41785 standard; DNA; 1266 BP.
XX
AC AAC41785;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33130.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 23-MAR-1999; 99US-0123548.
PR 25-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 99US-0136021.

```

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0155458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

seq_name:	/SID32/gcgdata/geneseq/NA2000.DAT:AAA81468
seq_documentation_block:	
ID	AAA81468 standard; DNA; 72750 BP.
XX	
XX	AAA81468;
XX	
XX	
DT	04-DEC-2000 (first entry)
XX	
XX	
DE	N. meningitidis partial DNA sequence gnm_16
XX	
XX	
KW	Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogen; vaccine; diagnosis; infection; antibacterial; antigen; vaccine; diagnosis; infection; antibacterial; immun

[illegible]

```

48860 AATTCCTGGAGCTGCTACTCTCCGGTTGCCACCTAAACAGAGGCTG 48811
190 euleutyrglylleleuGlnLeuGlnArgLysIleLysArgGlu 204
48810 TTATAGATGCCCTAACAAACACTCGTGAAGATATCGGAGAA 48767

```

seq_name: /SIDS2/gcgcdata/geneseq/geneseq/NA2000.DAT: AAC95299

seq_documentation_block:
ID AAC95299 standard; cDNA; 446 BP.

AC AAC95299;

DT 19-FEB-2001 (first entry)

DE Cat flea head and nerve cord (HNC) cDNA, SEQ ID NO:1794.

KW Cat flea; head and nerve cord nucleic acid; HNC; flea infestation;
KW vaccine; antiparasitic; therapeutic target; diagnosis; detection; ss.
OS Ctenocephalides felis.

PN WO200061621-A2.

PD 19-OCT-2000.

PF 07-APR-2000; 2000WO-US09437.

XX 09-APR-1999; 99US-0128704.

PA (HESK-) HESKA CORP.

PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

DR WPI; 2000-656323/63.

PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT acids useful for the prevention, diagnosis and treatment of flea
PT infestations -

PS Claim 26; Page 848; 964pp; English.

XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The
CC anti-HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HNC cDNA of the invention.

XX Sequence 446 BP; 179 A; 52 C; 75 G; 139 T; 1 other;

alignment_scores:
Quality: 418.00 Length: 84
Ratio: 5.160 Gaps: 0
Percent Similarity: 96.429 Percent Identity: 86.905
alignment_block:
US-09-525-867-1 x AAC95299
Align seg 1/1 to: AAC95299 from: 1 to: 446
130 MetAlaProAlaLeuArgLysValTyrAspGlnMetProGluProArgTyr 146
|||||
2 ATGGCTCCAGCGCTCGAAAAGTTTATGATCAATGGAGGAGTTATTACCATTTAT 51
146 rValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrS 163
|||||
52 GGTATCTCTATGGTAGTTGTGCTAATGGAGGAGTTATTACCATTTAT 101
163 eTyrSerValValArgGlyCysAspArgLysLeuValProValAspIleTyr 179
|||||
102 CATATCTCTAGTAGGGGTTGTGATAGAAATGATACCACTTGACATATAT 151
180 IleProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuG 196
|||||
152 GTACCAAGGATGCTCCAACTGCAGAAAGCTTTGTTATATGTTGTTTACA 201
196 nLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleTyrTyrArgA 213
|||||
202 GCTTCAAAAGAAAGTTAAACGAATGAAACCTGCAATGTTGGTATAGAA 251
213 rg 213
252 AA 253

seq_name: /SIDS2/gcgcdata/geneseq/geneseq/NA1998.DAT: AAV61275

seq_documentation_block:

ID AAV61275 standard; cDNA; 789 BP.

AC AAV61275;

DT 06-JAN-1999 (first entry)

DE 3' cDNA sequence of prostate tumour clone J1-21.

XX Prostate; cancer; tumour; vaccine; immunogen; clone; ss.

OS Homo sapiens.

PN WO9837093-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US03492.

XX 09-FEB-1998; 98US-0020956.

XX 25-FEB-1997; 97US-0806099.

XX 01-AUG-1997; 97US-0904804.

PA (CORI-) CORIAX CORP.

XX Dillon DC, Xu J;

XX WPI; 1998-609886/51.

XX Polypeptides comprising immunogenic portions of prostate proteins -

PT used in a vaccine for the treatment of prostate cancer

PS Claim 12; Page 48; 130pp; English.

XX The present sequence is a DNA which encodes an immunogenic portion

CC of a prostate tumour protein. The encoded immunogen, or the DNA itself,

CC can be used as a vaccine for the treatment of prostate cancer. The DNA

CC was identified by analysis of a subtracted cDNA library obtained by
 CC subtracting a prostate tumour cDNA expression library with a normal
 CC tissue cDNA library.

XX Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

alignment_scores:
 Quality: 411.00 Length: 189
 Ratio: 3.067 Gaps: 5
 Percent Similarity: 70.899 Percent Identity: 58.201

alignment_block:

US-09-525-867-1 x AAV61275/rev ..

Align seg 1/1 to reverse of: AAV61275 from: 1 to: 789

```

36 ThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVala 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
696 ACCGGGGCCCGAAGGGGAGCCAGCCAGCTGCGCNCACCAAGNAGAGCGGTG 647

52 AlaProLysProSerSerArgGlyGluTyrValValAlaLysLeuAspA 69
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
646 GGTTCCTCAACCCAGAACCGGGGGAGGATTGGGGGCCCAAGGGGATG 597

69 sp.LeuValAsnTrpAlaArgSerSerLeuTrpPro..... 81
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
596 ACTTGTGTG.....AACTGGNCCCGCGAGNTTT 565

82 .....MetThrPheGlyLeuAlaCysCysAlaValGluMetM 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
564 TTGTGGGCCCNACAAAGACTTTTGGGTGNCCTGTNGCCCGGAGGTGA 515

94 etHisMetAlaAla.ProArgTyrAspMetAspArg.PheGlyValValp 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
514 TGNANATNGCAGCACCCGNNAGACATGNCCTGTTTGGGNGNTTT 465

110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464 TTGGGGCCAGCCCGGCGCAGCTGTGAGTATGATTTGGCNGGNANATA 415

127 ThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
414 ACCAAAGAGAGGCCCGCCAGNCTTTGC.AAGTTTAAAGACCATGCGGA 366

143 uProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyrT 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
365 GCCCGGANAAGTGTTCCTCCANGGGGAGTGN.GCCACCGNGNGGTAAN 317

160 yfHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVal 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316 ACCACTATTCCTAATNGTGTGTGAGGNGTGCACCGCATTTGTGCCGTG 267

177 AspIleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG 193
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 GAAATTNAATCCAGGTGCGCCACCNANGCCGAGGCCCTGTTTANGG 217

193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 CATCTGTCAGTGTGAGAGGAANATCAAGCCGGAGCGGAGGCTGCANATT 167

210 rPtyrArgArg 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 GTTACCGCAGG 156

```

seq_name: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV58509

seq_documentation_block:

ID AAV58509 standard; cDNA; 789 BP.

XX

AC

XX

DT 08-DEC-1998 (first entry)

```

XX 3' fragment of prostate tumour specific gene J1-21.
DE Prostate tumour specific gene; human; prostate cancer; detection;
XX therapy; ss.
KW Homo sapiens.
XX WO9837418-A2.
XX 27-AUG-1998.
XX 25-FEB-1998; 98WO-US03690.
XX 09-FEB-1998; 98US-0904809.
XX 25-FEB-1997; 97US-0806596.
XX 01-AUG-1997; 97US-0904809.
XX (CORI-) CORIXA CORP.
XX Dillon DC, Xu J;
XX WPI; 1998-480805/41.
XX Novel human prostate specific tumour protein and fragments - useful
PT for detecting and treating prostate cancers
XX Claim 1; Page 50; 141pp; English.
XX This sequence represents a human prostate tumour specific gene, and can
CC be used in the method of the invention. The method is for detecting
CC prostate cancer comprising contacting a biological sample with an agent
CC able to bind an immunogenic portion of a prostate protein (such as
CC encoded by this sequence). An antibody which binds to an immunogenic
CC portion of the prostate protein, and the method can be used to detect,
CC monitor progression of, or treat prostate cancers. The antibody may
CC also be conjugated to a therapeutic agent for use in therapy of prostate
CC cancers.
XX Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;
SQ

```

alignment_scores:

Quality: 411.00 Length: 189
 Ratio: 3.067 Gaps: 5
 Percent Similarity: 70.899 Percent Identity: 58.201

alignment_block:

US-09-525-867-1 x AAV58509/rev ..

Align seg 1/1 to reverse of: AAV58509 from: 1 to: 789

```

36 ThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVala 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
696 ACCGGGGCCCGAAGGGGAGCCAGCCAGCTGCGCNCACCAAGNAGAGCGGTG 647

52 AlaProLysProSerSerArgGlyGluTyrValValAlaLysLeuAspA 69
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
646 GGTTCCTCAACCCAGAACCGGGGGAGGATTGGGGGCCCAAGGGGATG 597

69 sp.LeuValAsnTrpAlaArgSerSerLeuTrpPro..... 81
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
596 ACTTGTGTG.....AACTGGNCCCGCGAGNTTT 565

82 .....MetThrPheGlyLeuAlaCysCysAlaValGluMetM 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
564 TTGTGGGCCCNACAAAGACTTTTGGGTGNCCTGTNGCCCGGAGGTGA 515

94 etHisMetAlaAla.ProArgTyrAspMetAspArg.PheGlyValValp 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
514 TGNANATNGCAGCACCCGNNAGACATGNCCTGTTTGGGNGNTTT 465

110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126

```

```

|| .....T||
464 TTGGGCGCGCCCGCCAGTCTCAGTATGATGTGCGNGGNANNTA 415
127 ThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG1 143
||:||||| ||||:||||| ||||| ||||| ||||| |||||
414 ACCAAAAGAGAGGCCCGCCAGNGTTGC.AAGGTTTAAGACCAGATGCCGGA 366
143 uProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyrT 160
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 GCCGCGANAGTGGTTCCANGGGGAGTGN.GCCAAACGGNGGGGTAAN 317
160 yHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVal 176
||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
316 ACCACTATTCTTAATNGTGTGTGAGGGGNTGGACCCGATTGTGCCCGTG 267
177 AspIleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG1 193
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
266 GAAATNAAATCCAGGGTGCCACCNANGGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 CATCTGCAGCTGCAGAGGAANATCAACGGGAGCGGAGGCTGCANATT 167
210 rpTyrArgArg 213
|||||
166 GGTACCGCAGG 156
```



```
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS Mao,X.M., Xie,Y., Huang,X.Y., Ying,K. and Dai,J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China

FEATURES             Location/Qualifiers
     source            1..754
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="016403"
                        /tissue_type="brain"
                        /dev_stage="fetus"
                        10..654
     CDS               120 a 267 c 239 g 128 t
     ORIGIN
1 ccaagatggcgggtgctcagctccctggctgcggcttcggagatcctgtgtgcgct 133
2 |
3 |
4 |
5 |
6 |
7 |
8 |
9 |
10 |
11 |
12 |
13 |
14 |
15 |
16 |
17 |
18 |
19 |
20 |
21 |
22 |
23 |
24 |
25 |
26 |
27 |
28 |
29 |
30 |
31 |
32 |
33 |
34 |
35 |
36 |
37 |
38 |
39 |
40 |
41 |
42 |
43 |
44 |
45 |
46 |
47 |
48 |
49 |
50 |
51 |
52 |
53 |
54 |
55 |
56 |
57 |
58 |
59 |
60 |
61 |
62 |
63 |
64 |
65 |
66 |
67 |
68 |
69 |
70 |
71 |
72 |
73 |
74 |
75 |
76 |
77 |
78 |
79 |
80 |
81 |
82 |
83 |
84 |
85 |
86 |
87 |
88 |
89 |
90 |
91 |
92 |
93 |
94 |
95 |
96 |
97 |
98 |
99 |
100 |
101 |
102 |
103 |
104 |
105 |
106 |
107 |
108 |
109 |
110 |
111 |
112 |
113 |
114 |
115 |
116 |
117 |
118 |
119 |
120 |
121 |
122 |
123 |
124 |
125 |
126 |
127 |
128 |
129 |
130 |
131 |
132 |
133 |
134 |
135 |
136 |
137 |
138 |
139 |
140 |
141 |
142 |
143 |
144 |
145 |
146 |
147 |
148 |
149 |
150 |
151 |
152 |
153 |
154 |
155 |
156 |
157 |
158 |
159 |
160 |
161 |
162 |
163 |
164 |
165 |
166 |
167 |
168 |
169 |
170 |
171 |
172 |
173 |
174 |
175 |
176 |
177 |
178 |
179 |
180 |
181 |
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 |
782 |
783 |
784 |
785 |
786 |
787 |
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 |
842 |
843 |
844 |
845 |
846 |
847 |
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 |
902 |
903 |
904 |
905 |
906 |
907 |
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 |
962 |
963 |
964 |
965 |
966 |
967 |
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |
```

RESULT 3
AF060512

Best Local Similarity 69.8%; Pred. No. 1.9e-41;
 Matches 416; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 127 ctgcgctcagagtggtggccgcgtgtagcagcagaggtgtccatcagagcgtggccacc 186
 |||||
 Db 22093 CNGCTGTCAGTGGAGTGTCTGCATCCAGTGGTTCGTGCCAGAGACCATGCGGGT 22152
 |||||
 QY 187 gatggcccaagcagcagccagccctgcccgaagggccagagccgtgggtcccaacc 246
 |||||
 Db 22153 TACGATTACCGTGTCCAAATCCGCCAAAGTGGGCTACTCGCGCTCGGACCAACCCAG 22212
 |||||
 QY 247 agcaccggggcagtagtgtgtggcgaactgatacactctcaactgagcccccgg 306
 |||||
 Db 22213 AGACCTGGGGAGTGCACCTGTGCAGGTGTGATGACCTGTAACCTGGGACCAAG 22272
 |||||
 QY 307 agttctgtggccatgacctctggcctggcctgctgcgcctggagatgatgcacatg 366
 |||||
 Db 22273 GGATCCCTCTGGCGCTGACCTTTGGTCTCGCTGCTGCGCGCTGCGAAATGATGCACAT 22332
 |||||
 QY 367 gcagcaccgccctacgacatgagccgtttggcgtgtcttcgacccagccgcgcag 426
 |||||
 Db 22333 GCGGCTCTCGTACGACATGATGCTGATGCTGTGTGTTCCGAGCATCTCTCGCCAG 22392
 |||||
 QY 427 tcgcagctcatgctgtggcggcgacactcaccacaaagatggcccgagcgttcgcaag 486
 |||||
 Db 22393 GCGGAGTGTCTCATCTGCGCGGAACCTGACCAACAAGATGGCACCGGCTTTCGGAAG 22452
 |||||
 QY 487 gtctacgacagatgcggagccgcgtacgtgtctccatggggagctgcgcacagga 546
 |||||
 Db 22453 AFTCTACGACAGATGCGCGAGATGGGTGATTTCCATGGGCGATGTCGCCCAATGTT 22512
 |||||
 QY 547 ggaggtactaccacttctactcgtgtgtggtgggggtgcgacgcgtgtgcgcgtg 606
 |||||
 Db 22513 GCGGCTACTACCACTACTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 22572
 |||||
 QY 607 gacatctacatccagcgtgcccacccatgcggcggagggccctgtctacggatcctgcag 666
 |||||
 Db 22573 GACATCTACGTCGCCGGATGTCGCCACCGCCGAGGCTTAATGTACGGAATCTCTGCAG 22632
 |||||
 QY 667 ctgcagagaaatcaagcgggagcagagctacagatcgttgcacgaggtatgcg 722
 |||||
 Db 22633 CTGCAGAGAGAGTAAACCGCATGAGACCTGCAGATGTGGTACAGGAAGTAGAG 22688
 |||||

RESULT 9
 AE003770
 LOCUS AE003770 224100 bp DNA INV 05-OCT-2000
 DEFINITION Drosophila melanogaster genomic scaffold 14200001386035 section 95
 of 105, complete sequence.
 ACCESSION AE003770 AE002708
 VERSION AE003770.1 GI:7301790
 KEYWORDS HTG.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 224100)
 AUTHORS Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
 Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
 Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
 Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
 Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
 Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
 Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
 Benos,P.V., Berman,B.P., Bhandari,B., Bolshakov,S., Borkova,D.,
 Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
 Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
 Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
 Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
 Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,

Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S.,
 Fleischmann,W., Fowler,C., Gabriellian,A.E., Garg,N.S.,
 Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
 Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
 Wei,M.H., Ibegwam,C., Jaitani,M., Kalush,F., Karpen,G.H., Ke,Z.,
 Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
 Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
 Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
 McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
 Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
 Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
 Nusskern,D.R., Paclet,J.M., Palazzi,M., Picotman,G.S., Pan,S.,
 Pollard,J.V., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
 Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kianos,I.,
 Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
 Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
 Weinstein,G.M., Weissbach,J., Williams,S.M., Woodage,T.,
 Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
 Zaveri,J.S., Zhang,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
 Zhong,J.S., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
 Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
 The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)
 20196006
 2 (bases 1 to 224100)
 Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.
 Direct Submission
 Submitted (21-NOV-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA

FEATURES
 Location/Qualifiers
 1..224100
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="3R"
 <6113..>6751
 /gene="CG2014"
 /product="CT6471"
 /db_xref="FLYBASE:FBan0002014"
 /db_xref="FLYBASE:FBgn0039669"
 <6113..>6751
 /gene="CG2014"
 /map="99B3-99B4"
 /db_xref="FLYBASE:FBan0002014"
 /db_xref="FLYBASE:FBgn0039669"
 6113..6751
 /gene="CG2014"
 /note="CG2014 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0002014"
 /db_xref="FLYBASE:FBgn0039669"
 /protein_id="AAF56901.1"
 /db_xref="GI:7301791"
 /translation="MISFLKLRNPQQLLPSSGVLSQWLKQKTMPIVYRCONPP
 KWGYSAFRNQKTGWETCARDDLNNWGRKSLMPLTEGLACCAVEMHHAAPRYDM
 DRYGVFRASPRQADLIYVAGLTNNKAPAFKIXYDQMPERWIVSMSCANGGGYYH
 YSIVVRGCDRIVPDIVYVGGPPTAEALMYGILQKKKVRMTLQWYRK"
 join(46705..48163,54471..55572)
 /gene="msh"
 /note="Nucleotide sequence of the Celera sequence differs
 from the published sequence of this transcript."
 /product="CT2801"
 /db_xref="FLYBASE:FBan0001897"
 /db_xref="FLYBASE:FBgn0002849"
 <46705..>55572
 /gene="msh"
 /note="CG1897"
 /map="99B4-99B5"
 /db_xref="FLYBASE:FBan0001897"
 /db_xref="FLYBASE:FBgn0002849"
 join(47104..48163,54471..54808)
 /gene="msh"

source
 mRNA
 gene
 CDS
 mRNA
 gene
 CDS


```

CDS
1..333
/partial
/gene="uvrA"
/function="ATPase and DNA binding protein"
/notes="putative"
/codon_start=1
/transl_table=11
/product="DNA repair protein"
/protein_id="AAA03034.1"
/db_xref="GI:150619"
/translation="KLSKESSRATKTYLIIDPTGLGHFEDVRKLEVLHSLVQDQ
NTWVIEHNLVDVINTADWIIDIGPEGDGGGAIVATGTPQOVAEVPESHTRYLAPL
DAARSAAR"
880..1245
/gene="NQ07"
880..1245
/gene="NQ07"
/standard_name="nd3 subunit of Paracoccus NDH-1"
/EC_number="1.6.5.3"
/notes="GTG start codon; putative"
/codon_start=1
/transl_except="(pos:880..882,aa:Met)"
/transl_table=11
/product="NADH dehydrogenase"
/protein_id="AAA03035.1"
/db_xref="GI:150620"
/translation="MEVLLQYLPILVFLGMSALAIVLILAAVIAVRNPDPKYSV
YECGNAYDDARMKFDVRFVLVLSILILFLEVAFLFRWVSFASLSDVAFWGLKWL
AVLTGFAVEKKGALEWA"
1245..1769
/gene="NQ06"
1245..1769
/gene="NQ06"
/standard_name="ndhk subunit of Paracoccus NDH-1"
/EC_number="1.6.5.3"
/notes="putative"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase"
/protein_id="AAA03036.1"
/db_xref="GI:409251"
/translation="WTGINTAGADRLATAELNRELQDKGELLTTTEDIINWARGSL
HWMTPLGACCAVEMQIISMPRYDLERFGTAPRASPRQSDLMIVAGTLINKMAPALRY
YQWPEPRYISWMSGCANGGYTHYSYVWGCDRIVPDIIYVPGCPPTAELLIGLI
QIQRIRRTGTLVR"
1773..2399
/gene="NQ05"
1773..2399
/gene="NQ05"
/standard_name="25k subunit of Paracoccus NDH-1"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="NADH dehydrogenase"
/protein_id="AAA03037.1"
/db_xref="GI:409252"
/translation="MSALSDEALLLEAHLALRRENDVISTQVAFGELTVNATLSGV
IGLIFLNDPNCRFSLIDITADNPAPARFDVYHLLSKYQNRIRYKQVQRDE
LWPSLIGTGPQANNVREVEFDLFGISGDLRIIDYGFGRGHLRKDFPTTGVGE
VWSIDIEKRVYVYVNLVQVRYQFDLSFWEGAKYVLFQDKEAPEAK"
2453..3691
/gene="NQ04"
2453..3691
/gene="NQ04"
/standard_name="48k subunit of Paracoccus NDH-1"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="NADH dehydrogenase"
/protein_id="AAA03038.1"
/db_xref="GI:150623"

```

```
FEATURES             Location/Qualifiers
  source             1..29563
                    /organism="Drosophila melanogaster"
                    /db_xref="taxon:7227"
BASE COUNT          8656 a 5105 c 6205 g 8597 t
ORIGIN
133 tccagctggggccggctgtgcagccacagctggttcacatcagagctggccacccagatggc 192
19617 TTGGCCCTGCCCGCTGAAATTTGGCCCTGGTGCCCAACCAACAAACGCTTCCAGTTGC 19676
193 ccaagcagcaccagcctgcctgcgcaaaagccagagcctggtctcccaaaccccgagc 252
19677 CGAAGTGGCCCAAAATGCCCCAAGAAGGCTACTCTCGTTCCGGCCACCAACACATCTCTC 19736
253 cggggc-gagatgtggtgccaagctgagatgacctcgctcaactggccggcggagttc 311
19737 CGTTGCGGAGTGTCTACTGTGCCAGACTGGACGATCTGCTCAACTGGGTGCGCAAGGGCTC 19796
312 tcctgtgcccatacctctggcctgcccctgctgcctgcctggagatgacatggcagc 371
19797 GAICTGGCCACTGACTTTCGGTTGGCCCTGCTGTGCGGTCGAAATGATGACATCGCTGC 19856
372 acccgcacagatgagacgcttggctggtgtcttccgcgcagccgcgcgcagctcga 431
19857 TCCCGTTAGACATGATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 19916
432 cgtcatatgtggcggcgcacatcccaacaagatggcccgagcgtctcgcaagttcta 491
19917 TGTATCATGCTGCTGGCAGCGTGCACCAACAAATATGGACCGGCCCTCGCAAAAGTCTA 19976
492 cgaccagatcccgagcgcgctacgtggtctccatggggagctgcgcacacgggggag 551
19977 CGACCAATGCCGAGCCAGCGTTGGTGGTCTATCTCCATGGCGACGTGTGCCAACGGCGGG 20036
552 ctactacacattctactcgtggtggtgaggggctgcgacccatcgtgcccgtggacat 611
20037 CTACTACATTACTCTGACTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20096
612 ctatctcccggtgcccactcagccgagccctgctctcagcgcattcctcagctcga 671
20097 ATAGTACCCGGTGTGTCGCCCAACCCCGAGGCGCTCATGTAGCGGCTTTCACACTGCA 20156
672 gaggaagatcaagcggagcggaggtgcagatctgtgtaccgcaggtagcgc 723
20157 GAAGAAGTTAAGCGATGAAGACGCTCCAGATGTGGTATAGGAAGTAAAGC 20208

RESULT 13
AC010919            140685 bp      DNA              13-DEC-1999
LOCUS              Drosophila melanogaster chromosome x clone BACR32K23 (D891) RPCI-98
DEFINITION         32.K.23 map 14A-14A strain y; cn bw sp, *** SEQUENCING IN PROGRESS
                    ***, 77 unordered pieces.
AC010919            9 GI:6563419
VERSION            AC010919.9
KEYWORDS            HTG; HTGS_PHASE1.
SOURCE              fruit fly.
ORGANISM            Drosophila melanogaster
                    Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
                    Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
                    Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 140685)
AUTHORS            Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
                    Butenhofer,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
                    Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
                    Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
                    Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Pacleb,J.M., Park,S.,
                    Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Shih,E.,
                    Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
                    Rubin,G.M.
                    Submitted (27-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
                    Laboratory, MS 64-121, Berkeley, CA 94720, USA
                    On Dec 13, 1999 this sequence version replaced gi:6449494.
                    For further information about this sequence, including its location
                    and relationship to other sequences, please visit our sequence
                    archive web site (http://www.fruitfly.org/sequence/) or send email
                    to bgp@fruitfly.berkeley.edu. All contigs in this submission meet
                    the following cutoffs: length >= 200 bases.
                    * NOTE: This is a 'working draft' sequence. It currently
                    * consists of 77 contigs. The true order of the pieces
                    * is not known and their order in this sequence record is
                    * arbitrary. Gaps between the contigs are represented as
                    * runs of N, but the exact sizes of the gaps are unknown.
                    * This record will be updated with the finished sequence
                    * as soon as it is available and the accession number will
                    * be preserved.
                    * 1 674: contig of 674 bp in length
                    * 675 754: gap of unknown length
                    * 755 1507: contig of 753 bp in length
                    * 1508 1897: gap of unknown length
                    * 1898 2889: contig of 1302 bp in length
                    * 2890 2969: gap of unknown length
                    * 2970 3554: contig of 585 bp in length
                    * 3555 3634: gap of unknown length
                    * 3635 4240: contig of 605 bp in length
                    * 4240 4320: gap of unknown length
                    * 4320 4699: contig of 379 bp in length
                    * 4699 4778: gap of unknown length
                    * 4778 5295: contig of 517 bp in length
                    * 5295 5376: gap of unknown length
                    * 5376 5935: contig of 560 bp in length
                    * 5935 6015: gap of unknown length
                    * 6015 6641: contig of 545 bp in length
                    * 6641 7461: contig of 820 bp in length
                    * 7461 9170: gap of unknown length
                    * 9170 9250: contig of 1630 bp in length
                    * 9250 9852: contig of 632 bp in length
                    * 9852 9983: gap of unknown length
                    * 9983 11247: contig of 1284 bp in length
                    * 11247 11326: gap of unknown length
                    * 11326 12802: contig of 1476 bp in length
                    * 12802 12883: gap of unknown length
                    * 12883 14230: contig of 1348 bp in length
                    * 14230 14310: gap of unknown length
                    * 14310 15711: contig of 1401 bp in length
                    * 15711 15712: gap of unknown length
                    * 15712 16237: contig of 446 bp in length
                    * 16237 16318: gap of unknown length
                    * 16318 17991: contig of 1674 bp in length
                    * 17991 18071: gap of unknown length
                    * 18071 19191: contig of 1120 bp in length
                    * 19191 19271: gap of unknown length
                    * 19271 21324: contig of 2053 bp in length
                    * 21324 21404: gap of unknown length
                    * 21404 23898: contig of 2494 bp in length
                    * 23898 2405
```


Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacle,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Swirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstein,G.M., Weissenbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)

20196006

2 (bases 1 to 327446)

Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.

Direct Submission

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

On Oct 9, 2000 this sequence version replaced gi:7293108.

Location/Qualifiers

1..327446

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/chromosome="X"

complement(join(772..903,968..1504,1572..2077,2153..2821,4145..>4380))

/gene="CG9281"

/product="CT26402"

/db_xref="FLYBASE:FBan0009281"

/db_xref="FLYBASE:FBgn0030672"

complement(join(772..903,968..1504,1572..2077,2153..2813))

/gene="CG9281"

/note="CG9281 gene product"

/codon_start=1

/db_xref="FLYBASE:FBan0009281"

/db_xref="FLYBASE:FBgn0030672"

/protein_id="AAF48493.1"

/db_xref="GI:7293109"

/translation="MSSAAKKREAQKKDKARKNKIQIPSTKKANGVPERLTEEKL
 CALEEARISAEARCTGSLAVHPSROVKIANFSITFGSELLQDTMLENCGRY
 GLTIGCGKSSLLAVLGGREVPPPHIDIFHLTRI PASSKALQCYMEVDEIRKL
 EKLAELAMSEEDAOEQLIDYVERLDDMSADLAEVKARIILHGLGFGKAMQOKAKD
 FSGWRNRILARALPVKPHLLLDDEPTNHLDDACVWLEELKYKILVLSHQD
 FLNGVCNIHLTKKRLKYTGVAEAFVRMELLENQMKQINWQDDQLSHKNRIAR
 FHGSKLAQAQSKETLAKVAQGLTEKVDSDVLNFPFSCGKVPVPMVQNVN
 FRYNDEPWTYKLEFGIDIDTRALVPGNGAGKSTLLKLYGLDLPVTSGMRKNSHL
 RIARYHOHLLEHLDASPLEYMWRAFPPVKEEMRKIGRYGLTGRCQVCPIRQLS
 DGQRCVVFAWLAWQVPHLLDEPTNHLDMETIDALADAINDFDGMVLVSHDFRLI
 NQVARELWCEKTYKWKGGILDYKDLHKNKITSENEKKAKAG"

join(<7954..5295,5348..>5848)

/gene="CG15601"

/product="CT35724"

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

/evidence=not_experimental

<4963..>5848

/gene="CG15601"

/map="13E14-13E14"

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

CDS

join(4963..5295,5348..>5848)

/gene="CG15601"

/note="CG15601 gene product"

/codon_start=1

/db_xref="FLYBASE:FBan0015601"

/db_xref="FLYBASE:FBgn0030673"

/evidence=not_experimental

/protein_id="AAF48494.1"

/db_xref="GI:7293110"

/translation="WSKLSDKTPIPTETLEAYRROPCLYNLLDSYKNRVSRERAYGA
 IYRSKIPQTVSDIKLKIKSVRTYSEKELRIWMREKELGRTIYEPKLEFRADSFRLR
 IYSLSKRQKQGNSSAQLTTIKSDTSLKLTAAADITMSDEADEEDAEVNGEPE
 ECPLEBSRTASICKDDSLCLADQOEHYSGQSSQQLPHTMAQRKSYITSLDS
 AGEDDLIFGQSIASQLRTIPDSYRSYAKLRIOQVLEAETGQGTQSTEVASTOLQNT
 F"

join(<7954..8220,8556..8795,8865..9086,9145..>9416,
 9497..9993,10131..10515,10567..15238,15758..18218,
 18282..18755,18814..19154,19217..20287,20378..20513,
 20573..22032,22060..23197,23270..23671,23739..23927,
 23992..24313,24377..24542,24617..24807,24885..>24987)

/gene="CG8184"

/product="CT24358"

/db_xref="FLYBASE:FBan0008184"

/db_xref="FLYBASE:FBgn0030674"

/evidence=not_experimental

/protein_id="AAF48495.1"

/db_xref="GI:7293111"

/translation="MKVDVAKLKGSSELTAECLRIDELTKRRTRAELEFLDTVHV
 WIYKGELEHWEILDRFVKLLIEASRHYNGEFLVOCDFQETNLTLLHVLNFTT
 LLIEHFSRHLNYSIEHLIVLSEASQNDIVLAVLLLYMFSKSNFIPLRPEKKELL
 IVKLFNIAERWDPNYGLSKDCGCEKLEFLYQLCIDYVDYDHERHAQLEIFDMMDL
 CHTASAEVTKIAGQISKPSAKMIRIAHRLISGFNNYKLRLOVQARQVAVSIL
 IYSNALQDNTKVLVAGFGECELDKEDVHVEIRAVLRTLSMLHDFDNWVPS
 RAGSLMKIVHYTGAEQGTPLLVDCIDNLTTHGTERYPLIATSLFLSLYHLA
 SYWGCTALVKSQMOSLLCVISWPGVDLEHITFVTRAVRVIDLTINIDIAFHQNG
 LNVFDRLEIKSCAKALADIGITLKESTLRDDHKLDTSLDHVEDVEDGSGENASV
 TSGNASPRDNDSDGATGAGSIVYRAYHEHIGAGSGAGLLDIGARSLSGSGSTS
 GMGASINRDGSAARPPAIMYPTNYYIARPLERKAELSLSPLSLAKPKPCVQRA
 ALLNNRNIMETSLTQSLRHIANAEYYPGLFATDVTVTVVNEPSLLSSLODLG
 ITSMVKALLQKQVPATREVIGSLPNVFSALCINERGLFELSYDPFKVLKVLSPD
 YLVAMRRRSDDPLGDTATNLGNMDDLKHHPYLRADATEAIRLLNELVRLGSDPS
 FICWRANKESGSGSGSGHGHVASTPSPMVVAGTGSASSVLOGAADTNDSSGD
 DDDDDMSASQOQQPTPGCGGSPPTQOAGVVGSGAGATGTGTPAAASQAVKY
 VTPPERAFLIDILYNMKFIEAFISNPGDHCFREFVLHGLAPILQLLSLNLPLV
 DSVSTTSQAVANVCALVSAQETKVLDAKLDIADIVAQKLKIKHTIFPGGSVLL
 ETGVQLQOVLQVILSVWNETILLNCSDEPTQHPDLIDWEIAQMSAAGTADPLT
 EAEISRLERAARFRTYPTPEOFYIKSLAASSRLGRALAEGLLTVKLSVSPQT
 RQRINDLISNITKVTPEARITILSYNGFESHILPKVPKLTFLICSVG
 FTGPMFLDKRKAFLHMSQFAENGKKAFFEMFYWALSIDNVSKPFEDQWASMEDL
 MQSHEDDKRCPGEGTFEDAWQLLEKAVNPRMLDSDPYTHSPALNYPDVSVPREPV
 FYLIHOLAMQALRIWARRPINYPGSPMETMILIKHLKSHQTLDRHTHRKE
 IKFENLYIRSTDGLNADHKTLDGMGFHYHVEIALRTNASEADYLLNPPASS
 LSTTGQSSAPPAPPPPPSTMDIDVDPADGESQSKSTSPNSDYKHLKMLP
 SLIFRCDDEAISRAREWEAIPDVTNSAADIAVLYRRHNLKVFYVNEVRLIQ
 WYFTQLPFPKATGCSKASLEECQLTGATFLFVRLKTLSTLLPENYSDLRPLVE
 ALTAQDAWVSLVKLLDCSQMWLEQSTENTGOORATVPRWQNLVDLDAIDSLSHI
 LQRKANRVSSDVKLLWYDASTGKNWAYSEANNELIRNAYAGERWLYINIGRQRTV
 SLNCMTCVSEASGTHRPVCPALKLSAITSNNPNVALHKVHILNRVVRTPDGSGTV

mRNA

gene

CDS

CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immunotherapy. A portion of
CC the polynucleotides encoding the polypeptides can also be used
CC to modulate the expression of the polypeptides. AA06221 to AA06691
CC AA92000 to AA92020 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 789 BP: 129 A; 240 C; 159 G; 195 T; 65 other:

Query Match	46.7%	Score 384.6;	DB 21;	Length 789;
Best Local Similarity	75.9%	Pred. No. 2.1e-75;		
Matches 527; Conservative	0;	Mismatches 154;	Indels 13;	Gaps 7;

QY	141	ggcccgctgtcaggcacaggttccatcagagctggccacgatggcccaagcag	200
Db	739	GGGGATGGTGNNAAGANGSNTTTCTNTTAGNNNGGCCACCGGGGGCCCGAAGG	680
QY	201	caccagcctgcctgtccaaaggccagagcctgtctccaaacccagcagccgagcga	260
Db	679	GGANCCANTTCCCNCCAAAGNNAGAGCGTGGGTTCCTAAACCCAGNANCGGGGGA	620
QY	261	gtatgtgtggccaagtgtgatcactcgt---caactggggccggcgagttctctgtg	317
Db	619	GGATTGGGGCCCAAGGGGATGACTTTGTAAACTGGNCCCGCCCGAGNTTTTTTGN	560
QY	318	g----cccatgacctcggcctgctgcgcgtgagatcatgacatgacagcagc-a	372
Db	559	GGCCNCAAGACTTTTGGGTGNCCTGNTGNCCTGGGAGGTGATGATNATNGCAGAA	500
QY	373	cccgtacgacatggaccgctttggcgtgg---tcttcgcgcagccgcgcagctccg	430
Db	499	CCCCNNAAAGACATGNCCTGGTTTGGGGNGNTTTTGGGCCAGCCGCGCAGTCTG	440
QY	431	acgtcataatcgtgcgcgcacactcaacaaagataggcccaagcttcgcaagctc	490
Db	439	A-GTNATGATTGTGCGNGNANATACCCAAAGAGGGCCCCAGNG-TTGTGAAGTTT	382
QY	491	acgaccagatccggagccgcgtacgtgctccatggagctgcgccaaaggagag	550
Db	381	AAGACCAGATCCGAGCCGACGANAAGTGGTTCCANGSGAG-TGNGCCACCGNGNG	323
QY	551	gtactacacatattcactcgtgtagtgtaggggtcgacgcgcgtcgcctgtgaca	610
Db	322	GGTAANACCATTATTCCTAATNGTGGTAGGGNTGGACCGATTGTCCCGTGGAAA	263
QY	611	tctacatccaggtgtcccaactacggccgagccctcgtcactcagcattcgcagctgc	670
Db	262	TTNAAATCCAGGTGCCCCANANGCCGAGGCCCTGTTTANGGCATCCTCAGCTGC	203
QY	671	agaggaagatcaaggaggagcgtgcagatcgtgtaccgcagtagcgcgcgcgc	730
Db	202	AGAGGAANATCAAGCGGAGCGGAGGCTGCANATTGTGTACCGAGGTAGGGCGCGCCG	143
QY	731	gcgcgcgcgagcctgtgcgcctcctgtcccaagcctgttgttccgttaagttgtc	790
Db	142	GCCGCGCGGAGCCTGTGCCCTCCTGTCCCCAGCCGNTGTGTGCCGNGAGGTTGTC	83
QY	791	aataacctgcctcgggcgaataaaaaa	824
Db	82	AATAAACCTGCCTNGNAAAAA	49

RESULT	7
AAS10031/C	
ID	AAS10031 standard; cDNA; 789 BP.
XX	
XX	
XX	AAS10031;
XX	
DT	24-OCT-2001 (first entry)
XX	
DE	Human prostate tumour cDNA J1-21
XX	

KW	Human; prostate tumour protein; prostate cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	US6262245-B1.
XX	
PD	17-JUL-2001.
XX	
PF	25-FEB-1998; 98US-0030607.
XX	
PR	25-FEB-1997; 97US-0806099.
PR	01-AUG-1997; 97US-0904804.
PR	09-FEB-1998; 98US-0020356.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Xu J, Dillon DC;
XX	
DR	WPI; 2001-440862/47.
XX	
PT	Novel polynucleotide encoding polypeptide comprising a portion of
PT	prostate tumour protein useful for inhibiting development of prostate
PT	cancer or for treating prostate cancer in a patient
XX	
PS	Example 1; Column 54-5; 105pp; English.
XX	
CC	The sequence is a human prostate tumour cDNA which encodes a
CC	partial tumour protein. The DNA is useful for inhibiting the development
CC	of prostate cancer or for treating prostate cancer in a patient.
XX	
SQ	Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

Query Match	46.7%	Score 384.6;	DB 22;	Length 789;
Best Local Similarity	75.9%	Pred. No. 2.1e-75;		
Matches 527;	Conservative	0;	Mismatches 154;	Indels 13;
				Gaps 7;

[illegible]

RESULT 11

KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0128264.
PR 29-MAR-1999; 99US-0128785.
XX 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0138817.
PR 22-JUN-1999; 99US-0138899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 09-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 11-JUN-1999; 99US-0139119.
PR 12-JUN-1999; 99US-0139452.
PR 13-JUN-1999; 99US-0139453.
PR 14-JUN-1999; 99US-0139454.
PR 15-JUN-1999; 99US-0139455.
PR 16-JUN-1999; 99US-0139456.
PR 17-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 19-JUN-1999; 99US-0139459.
PR 20-JUN-1999; 99US-0139460.
PR 21-JUN-1999; 99US-0139461.
PR 22-JUN-1999; 99US-0139462.
PR 23-JUN-1999; 99US-0139463.
PR 24-JUN-1999; 99US-0139464.
PR 25-JUN-1999; 99US-0139465.
PR 26-JUN-1999; 99US-0140823.
PR 27-JUN-1999; 99US-0140991.
PR 28-JUN-1999; 99US-0141287.
PR 29-JUN-1999; 99US-0141842.
PR 30-JUN-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 03-JUL-1999; 99US-0142803.
PR 04-JUL-1999; 99US-0142920.
PR 05-JUL-1999; 99US-0142977.
PR 06-JUL-1999; 99US-0143542.
PR 07-JUL-1999; 99US-0143624.
PR 08-JUL-1999; 99US-0144005.
PR 09-JUL-1999; 99US-0144085.
PR 10-JUL-1999; 99US-0144086.
PR 11-JUL-1999; 99US-0144325.
PR 12-JUL-1999; 99US-0144331.
PR 13-JUL-1999; 99US-0144332.
PR 14-JUL-1999; 99US-0144333.
PR 15-JUL-1999; 99US-0144334.
PR 16-JUL-1999; 99US-0144335.
PR 17-JUL-1999; 99US-0144352.
PR 18-JUL-1999; 99US-0144632.
PR 19-JUL-1999; 99US-0144684.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 23-JUL-1999; 99US-0145088.
PR 24-JUL-1999; 99US-0145089.
PR 25-JUL-1999; 99US-0145192.
PR 26-JUL-1999; 99US-0145145.
PR 27-JUL-1999; 99US-0145218.
PR 28-JUL-1999; 99US-0145224.
PR 29-JUL-1999; 99US-0145276.
PR 30-JUL-1999; 99US-0145313.
PR 31-JUL-1999; 99US-0145318.
PR 01-AUG-1999; 99US-0145319.
PR 02-AUG-1999; 99US-0145351.
PR 03-AUG-1999; 99US-0146386.
PR 04-AUG-1999; 99US-0146388.
PR 05-AUG-1999; 99US-0146389.
PR 06-AUG-1999; 99US-0147038.
PR 07-AUG-1999; 99US-0147204.
PR 08-AUG-1999; 99US-0147302.
PR 09-AUG-1999; 99US-0147302.
PR 10-AUG-1999; 99US-0147360.
PR 11-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 07-AUG-1999; 99US-0147493.
PR 08-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 14-AUG-1999; 99US-0149368.
PR 15-AUG-1999; 99US-0149175.
PR 16-AUG-1999; 99US-0149426.
PR 17-AUG-1999; 99US-0149722.
PR 18-AUG-1999; 99US-0149723.
PR 19-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149902.
PR 21-AUG-1999; 99US-0149930.
PR 22-AUG-1999; 99US-0150566.
PR 23-AUG-1999; 99US-0150884.
PR 24-AUG-1999; 99US-0151065.
PR 25-AUG-1999; 99US-0151066.
PR 26-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 28-AUG-1999; 99US-0151438.
PR 29-AUG-1999; 99US-0151930.
PR 30-AUG-1999; 99US-0152363.
PR 31-AUG-1999; 99US-0153070.
PR 01-SEP-1999; 99US-0153758.
PR 02-SEP-1999; 99US-0154018.
PR 03-SEP-1999; 99US-0154039.
PR 04-SEP-1999; 99US-0154779.
PR 05-SEP-1999; 99US-0155139.
PR 06-SEP-1999; 99US-0155486.
PR 07-SEP-1999; 99US-0155659.
PR 08-SEP-1999; 99US-0156458.
PR 09-SEP-1999; 99US-0156596.
PR 10-SEP-1999; 99US-0157117.
PR 11-SEP-1999; 99US-0157753.
PR 12-SEP-1999; 99US-0157865.
PR 13-SEP-1999; 99US-0158029.
PR 14-SEP-1999; 99US-0158232.
PR 15-SEP-1999; 99US-0158369.
PR 16-SEP-1999; 99US-0158929.
PR 17-SEP-1999; 99US-0159294.
PR 18-SEP-1999; 99US-0159295.
PR 19-SEP-1999; 99US-0159329.
PR 20-SEP-1999; 99US-0159330.
PR 21-SEP-1999; 99US-0159331.
PR 22-SEP-1999; 99US-0159637.
PR 23-SEP-1999; 99US-0159638.
PR 24-SEP-1999; 99US-0159584.
PR 25-SEP-1999; 99US-0160741.
PR 26-SEP-1999; 99US-0160767.
PR 27-SEP-1999; 99US-0160768.
PR 28-SEP-1999; 99US-0160770.
PR 29-SEP-1999; 99US-0160814.
PR 30-SEP-1999; 99US-0160815.
PR 01-OCT-1999; 99US-0160980.
PR 02-OCT-1999; 99US-0160981.
PR 03-OCT-1999; 99US-0160989.
PR 04-OCT-1999; 99US-0161404.
PR 05-OCT-1999; 99US-0161405.
PR 06-OCT-1999; 99US-0161406.
PR 07-OCT-1999; 99US-0161359.
PR 08-OCT-1999; 99US-0161360.
PR 09-OCT-1999; 99US-0161361.
PR 10-OCT-1999; 99US-0161320.
PR 11-OCT-1999; 99US-0161992.
PR 12-OCT-1999; 99US-0161993.
PR 13-OCT-1999; 99US-0162142.
PR 14-OCT-1999; 99US-0162143.
PR 15-OCT-1999; 99US-0162144.
PR 16-OCT-1999; 99US-0162145.
PR 17-OCT-1999; 99US-0162146.
PR 18-OCT-1999; 99US-0162147.
PR 19-OCT-1999; 99US-0162148.
PR 20-OCT-1999; 99US-0162149.
PR 21-OCT-1999; 99US-0162150.
PR 22-OCT-1999; 99US-0162151.
PR 23-OCT-1999; 99US-0162152.
PR 24-OCT-1999; 99US-0162153.
PR 25-OCT-1999; 99US-0162154.
PR 26-OCT-1999; 99US-0162155.
PR 27-OCT-1999; 99US-0162156.
PR 28-OCT-1999; 99US-0162157.
PR 29-OCT-1999; 99US-0162158.

Query Match 30.6%; Score 251.8; DB 21; Length 976;
Best Local Similarity 69.8%; Pred. No. 2.4e-46;
Matches 340; Conservative 0; Mismatches 147; Indels 0; Gaps 0;


```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
```

```
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      28.9%; Score 237.8; DB 21; Length 1266;
Best Local Similarity 69.4%; Pred. No. 2.8e-43;
Matches 338; Conservative 0; Mismatches 147; Indels 2; Gaps 1;

QY 201 caccagcctgcctgccaaaggccagagccgtggctcccaaacccagcagccggcgga 260
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 518 caccagacctgtctctctcgaactccctccctccctccctccctccctccctccctcc 577
QY 261 gtagtggtggccaaagctggatgacctcgtcaactcgtggcggcgaggtctctctgtgcc 320
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 578 atttgatctcccaag--gtatgatctcgaactggctggctggctggctggctggctggcc 635
QY 321 catgaccttcgctggcctgctgcgcctggagatgatgcacatggcagcccccctta 380
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 636 tatgaccttggctcgtctgctgcgcctggagaaatgatgcataccggtgctgctgcta 695
QY 381 cgaatggagccgcttggcgtggtcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 440
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 696 cgaatcgcagctcgtcgtatcattcctcagcctcagcctcagcctcagcctcagcctcag 755
QY 441 cgtggccgcgcacactcaccacaagatggcccccagccttcgaaaggtctacacacagat 500
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 756 tctgcgcgttactcttaccacaagatggctcccggtctcttcgcaaggttatgacacagat 815
QY 501 gccggagccgcgcgtacgtggtctccatggggagctgcgcacacagcagggaggtactacca 560
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 816 gctgagccaaagggtggtgatttcaatgggaagtgtgccaatggtggtgatactacca 875
QY 561 ctattcctactcgtggtgaggggctgcgacccgcatcgtcccggtgacatcacatccc 620
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 876 ctactcctactcgtggtgagggatgtgacagaatgtccacagtcacacatacagctccc 935
QY 621 aggtgcccacactacgcccagggcctgctctacggcatcctcagcagctgcagaggaagat 680
Dd      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dd 936 ggggtgcccacacacccgctgaggttggctctatggtactactccagcagcaggaat 995
QY 681 caagcgg 687
Dd      ||| ||
Dd 996 caacagg 1002
```

```
RESULT 15
AAA81468/c
ID AAA81468 standard; DNA; 72750 BP.
XX
AC AAA81468;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_16 SEQ ID NO:16.
XX
```


Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	715.8	85.9	774	11	BG706832	BG706832 602671984
2	706.8	85.8	835	11	BG709937	BG709937 602527493
3	702.4	85.2	799	10	AL580649	AL580649 AL580649
4	699.8	84.9	879	11	BE794266	BE794266 601591551
5	692.8	84.1	738	11	BG451103	BG451103 602199337
6	692.4	84.0	695	11	BG488658	BG488658 602534588
7	682.4	82.8	732	11	BI116730	BI116730 602868416
8	677.8	82.3	742	11	BG765790	BG765790 602739620
9	661.2	80.2	813	11	BG65478	BG65478 602738979
10	652.6	79.2	936	11	BF968344	BF968344 602369415
11	652.4	79.2	1032	11	BE795373	BE795373 601592903
12	651.8	79.1	1002	11	BG395342	BG395342 602457879

```

/lab_host="DH10B"
/ncce="Organ: Brain; Vector: pBluescriptR (modified
pBluescript Ks+); Site-1: BamHI; Site-2: SalI-XhoI (atccag
); Oligo-dn primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NIHGRI, National
Institutes of Health). Note: This is a NIH-MGC Library."
132 a 259 c 132 t
BASE COUNT

```

Query Match 86.9%; Score 715.8; DB 11; Length 774;
Best Local Similarity 98.0%; Pred. No. 1.4e-128;
Matches 746; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

[illegible]

RESULT 2
BG479937

LOCUS BG479937 836 bp mRNA 21-MAR-2001
DEFINITION 602527493f1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4651103 5',
mRNA sequence.
ACCESSION BG479937
VERSION BG479937.1 GI:13412216
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1433 row: i column: 24
High quality sequence stop: 751.
FEATURES
Location/Qualifiers
1..836
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4651103"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site.1: XhoI;
Site.2: EcoRI; CDNA made by oligo-3T priming
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
152 a 281 c 261 g 141 t 1 others

Query Match	85.8%	Score	706.8	DB 11	Length	836			
Best Local Similarity	98.5%	Pred. No.	7.5e-127						
Matches	744	Conservative	0	Mismatches	8	Indels	3	Gaps	3

QY	73	gcaaatatggcggctgctgaagctctctgacctgcggcggttcacgatactttggcttgagc	132
Db	1	GCCAAATGGCGGTGCTGTACATCTCTGGGCTCGGGGCTTCGGATCTGCTGTGGCG	60
QY	133	tccagcg-tggggcccgctgtgcaggcacaggtgtccatcacagcgtggccaccagtg	191
Db	61	TCCAGGCTGGCCCGGCTGTGCAGGCCACAGAGTGTCCATCAGAGCGTGCCACCGATGG	120
QY	192	ccaaagcagcacccagcctgcctgtccaaaggccagagcgtgtgtcccaaacccagcag	251
Db	121	GCCAAAGCACCCAGCGTGCCTTGCCAAAGGCCAGAGCGGTGGCTCCCAAAACCCAGCAG	180
QY	252	ccggggcgaagtatgtgtggccaagctggatgacctgtcaactggcccgccggagttc	311
Db	181	CCGGGCGGAGTATGTGTGGCCAAAGCTGGATGACCTGTCACTTGGGCGCCCGGAGTTC	240
QY	312	tctgtggccatgaaccttggccttggcctgtctgcgcgttggaatgatcacatggcagc	371
Db	241	TCGTGGCCCATGACCTTCGSCCTGGCCCTGCTGGCGCTGGAGATGATGCACATGGCAGC	300
QY	372	aacccgctacatggaacgctttggcgtgtgtttcccgccagcccgccagctccga	431
Db	301	ACCCGCTACGACATGAGACCGCTTTGGCTGTGTTCTCCGGGCGAGCCGCGGCACCGGA	360


```

Db 121 CCAAGCAGCACCAGCTGCGCCCTGCAAAAGCCAGAGCGGTGGTCCCAACCCAGCAGC 180
QY 253 cggggcgagtagtggtgccaagctgtagtaccctcgtcaactgggcccgcggagttct 312
Db 181 CGGGCGAGTAGTGTGGCCAAAGCTGGATGACCTCGTCACTGGGCCCGCGGAGTCT 240
QY 313 cgtgtggccatgacctcggcctggcctgctgctgcccgtggagatgacatggcagca 372
Db 241 CAGTGGCCATGACCTCGGCTGGCCCTGCTGCGCCGTGGAGATGATGCACATGGCAGCA 300
QY 373 ccccgtaagacatgagccgttttgctgcttccgcgcagccgcgcgcagctccgac 432
Db 301 CCCCCTAGCAGATGAGCCGCTTTGGCGTGTCTTCCGCGCCAGCCCGCGGAGTCCGAC 360
QY 433 gtcagtgtgtggccggaactcaccacaaagatggccccagcgcgttcgcaaggtctac 492
Db 361 GTCATGATCGTGGCCGGCACACTCACCACAAAGATGGCCCCAGCGCTTCGCAAGGTCTAC 420
QY 493 gaccagatgcggagccgcctagctgtctccatgggagctgcgcaacggagagc 552
Db 421 GACCCAGATGCCGAGCCCGCTACGTGTCTCCATGGGAGCTCGCCCAACGGAGAGGC 480
QY 553 tactaccactattcctactcgtgtggtgagggctgcgacgcacatcgtgccgtgacatc 612
Db 481 TACTACCACATATCTACTCGTGTGTGAGGGCTGCGACGCGCATCGTGGCCGTGGACATC 540
QY 613 tacatccaggtgcgccacctacggccgagggccctgtctctacgcatcctcgcagctgacg 672
Db 541 TACATCCCAAGGCTG-CCACCTACGGCCGAGGCGCTGTCTACGGCATCTCTGCAAGTCCAG 599
QY 673 aggaagatcaagcggagcggaggtgcagatctgtgtaccgcaggtgagccgcgcgcgc 732
Db 600 AGGAGATCAAGCGGAGCGGAGCTGCAGATCTGGTA-CCGAGTACGCGCCGCC- 656
QY 733 cggccggagcgtgtgcgcctcctgtcccgagcctgtgttcccgtaggttgcac 792
Db 657 CCGCGCCGAGGCTGTGCGCTCTGCTGCTCCCGAGCTGCTGTGCTGCTGCTGCTGCTGCTGCT 716
QY 793 taacactgccccgggcaaaa 814
Db 717 TAACCTGCCCCTCGGGCAAAA 738

RESULT 6
LOCUS BG488658
DEFINITION 602534588F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4662381 5',
mRNA sequence.
ACCESSION BG488658
VERSION BG488658.1 GI:13450165
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 695)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1462 row: o column: 22
High quality sequence stop: 691.
Location/Qualifiers
1..695

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4662381"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pORF7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

BASE COUNT 114 a 242 c 225 g 114 t
ORIGIN

Query Match 84.0%; Score 692.4; DB 11; Length 695;
Best Local Similarity 99.9%; Pred. No. 4.6e-124;
Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 70 gagggcaagatggcggtgctgtcagctcctgctgcggtcctccggatcccttggtctg 129
Db 2 GAGGCCAAGATGGCGGTGCTGTCAGCTCCTGCGCTCGCGGCTTCGGATCCTTGGTCTG 61
QY 130 cgtccagcgtgggcccggctgtgcaggcacgaggtgtccatcagagcgtggccacgat 189
Db 62 CGCTCCACGCTGGGCTGGCTGTGCAGGACAGAGGTGTCCATCAGAGCGTGGCCACGAT 121
QY 190 gcccgaagcagcaccagcctgcctgcctgcgaagccagagcgtggctcccaaccagc 249
Db 122 GCGCCCAAGCAGCACCAGCTGCCCTGCCAAGGCCAGAGCCGTGGCTTCCAAACCCAGC 181
QY 250 agcggggcgagtagtggtggccaagctggtgacacctgctcaactgggcccgcgcgaggt 309
Db 182 AGCCGGGCGAGTATGTGTGGCCAAAGCTGATGACCTCGTCAACTGGCGCGCGGAGT 241
QY 310 tctctgtggccatgacacttcggcctggcctgctgcgcctggagatgagcatggca 369
Db 242 TCTCTGTGGCCCATGACCTTGGCTGCGCTGCGCCGCTGGAGATGATGCATGGCA 301
QY 370 gcaacccgcctacgacatggacgcgttttggtggttccgcgcagccgcgcgcagtc 429
Db 302 GCACCCCGCTACGACATGACCGCTTTGGCGTGTCTTCGCGCCAGCCGCGCCAGTCC 361
QY 430 gactcatgactgtggccggaactcaccacaaagatggccccagcgcgttcgcaaggtc 489
Db 362 GACGTCTATGATCGTGGCGGCGACACTCACCACAAAGATGGCCCGCGCTTCGCAAGTTC 421
QY 490 tacgaccagatgcccgaagcgcgtacgtgtctccatgggagcgtgcgaacgagga 549
Db 422 TACGACCAAGTCCCGAGCCCGCTACGTGTCTCCATGGGAGCTGCGCCAAAGGAGGA 481
QY 550 ggctactaccactattcctactcgtgtggtgggggctgcgaccgcacatcgtgcccgtggac 609
Db 482 GGCTACTACCACTATTCTACTCGGTGTGGTGGGCTGCGACCGCATCGTGGCGTGGAC 541
QY 610 atctacatccaggtgcccactacggcgagccctgtctctacggcatcctgcagctg 669
Db 542 ATCTACATCCAGGCTGCCACCTACGCGCGAGCCCTGTCTACGGCATCTCTGCAAGTCTG 601
QY 670 cagaggaagatcaagcgggagcggagcgtgcagatctgttaccgcaggttagcgcgcgcgc 729
Db 602 CAGAGGAGATCAGCGGAGCGGAGGCTGCAGATCTGTGTACCGCAGGTAGCGCCGCCGC 661
QY 730 cgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 763
Db 662 CGCGCCGCGGAGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695

RESULT 7
B1116730

LOCUS B1116730 .733 bp mRNA EST 26-JUN-2001
 DEFINITION 602868415f1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5017179 5',
 mRNA sequence.
 ACCESSION B1116730
 VERSION B1116730.1 GI:14567631
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 733)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCMI828 row: 0 column: 04
 High quality sequence stop: 670.
 FEATURES
 Location/Qualifiers
 1..733
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5017179"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 130 a 233 c 230 g 120 t
 ORIGIN
 Query Match 82.8%; Score 682.4; DB 11; Length 733;
 Best Local Similarity 97.6%; Pred. No. 3.9e-122;
 Matches 703; Conservative 0; Mismatches 16; Indels 1; Gaps 1;
 QY 105 gcgcggttcggtacgttgcgtccagcgtggcgccggtgtgcaggcagcagg 164
 DB 2 gcgcggcttcggatccttgcgtccagcgtggcgccggtgtgcaggcagcagg 61
 QY 165 tgcctacagagcgtggccaccagtgcccaagcagcaccagcctgcctgccaaagc 224
 DB 62 tgcctacagagcgtggccaccagtgcccaagcagcaccagcctgcctgccaaagc 121
 QY 225 cagagccgtggtcccaaacccagcagcggggcgagtagtgcgtggccaaagcgtgatga 284
 DB 122 CAGAGCCGTGGCTCCCAAAACCACGACGCGGGCGAGTAGTGTGTCGCCAAGCTGATGA 181
 QY 285 cctgcctaactggccgcgcggagttctgtggcccaactgcctggcctggcctgctg 344
 DB 182 CCTGCTCAACTGGGCCCGCGGAGTTCTGTGGCCCATGACCTTCGGCTGGCCTGGT 241
 QY 345 qcgcgtgagatgatgcacatggcagcaccgcgtacagacatgagcgtttggcgtgct 404
 DB 242 CGCCGTGAGATGATGACATGGCAGCACCCCGCTACGACATGGACCGCTTGGCGTGT 301
 QY 405 ttctgcgcagccgcgcgcagtcgcagtcgtatgcgtggccggccacactcaccaaca 464
 DB 302 CTTCCGCCACGCGCGCGGACGTCGACGTCATGATCGTGGCGGCGACACTCACCAACA 361

QY 465 gatgccccagcgttcgcaaggtctacgaccagatgcggagccgcgtactagtgtc 524
 DB 362 GATGCCCCAGCGTTTCGCAAGGTTACGACCAGATGCCGAGCCGCTACTGTGTTCTC 421
 QY 525 catgggagctgcgccaacgagaggtactaccactattctactcgtggtgagggg 584
 DB 422 CATGGGAGCTGCCCAACGAGGAGGCTACTACCACTATTCTACTCGTGTGGTGGGG 481
 QY 585 ctgcgaccgcatgctgccctgacatctatccagagctgccacactacagcagcagc 644
 DB 482 CTGCAGCGCATGCTGCCCGTGACATCTACATCCAGGCTGCCACCTACGTGCCAGGC 541
 QY 645 cctgctcagcgcattcctgcagctgcagagagatcaagcgggagcggaggtcgcagat 704
 DB 542 CTTGCTCTACGGCATCTCTGCAGTGCAGAGAGAGATCAAGCGGAGCGAGGCTGCAGAT 601
 QY 705 ctgtacccagatagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 764
 DB 602 CTGTATCCGAGGTAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 QY 765 gctcgtgtgtccctgcaggtgtgtaataaacctgcctcgggcaaaaaa 824
 DB 661 GCCTGCTGTGTCGCTGAGTGTCTCAATAAAGCTGCTCGGCGGCTGCAGCAAGAA 720
 RESULT 8
 BG765790 742 bp mRNA EST 15-MAY-2001
 LOCUS 602739620F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4869532 5',
 mRNA sequence.
 ACCESSION BG765790
 VERSION BG765790.1 GI:14076443
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 742)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI741 row: 9 column: 05
 High quality sequence stop: 739.
 FEATURES
 Location/Qualifiers
 1..742
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4869532"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 155 a 245 c 224 g 118 t
 ORIGIN

Db 422 CGCAGGCTCTACGACAGATGCCGAGCCGCGCTACGTGCTCTCCATGGGAGCTGCCCC 481
 QY 541 aacgagagaggtactaccactattctactcgtggtgaggggctgagccagcagctg 600
 Db 482 AACGAGAGAGGCTACACACATCTTCTACTCGGTGTGAGGGGTGCGACCGCATCGT 541
 QY 501 ccgctggacattacattccaggctgccacactacagccagccccc-tgctetacggcat 659
 Db 542 CCGGTGACATCTACATCCCAAGGCTGCCACCTACGGCCGAGCGCCCTTCTCTACGGCAT 601
 QY 660 --cctgcagctgcagaggaagatcaagcggagcggaggtcagatctgttacccagcag 717
 Db 602 TCCTGAGGTTGACAGAGGATCAAGCGGAGCGGAGGAGCTCCAGATCTGTACCGCAGG 661
 QY 718 tagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 777
 Db 662 TAGCGCGG----CGCGCGCGCGGACCTGTGCGCGTCTCTGT-CCAGCGCTCTGTGTGTC 716
 QY 778 cctgaggttgtcaataaaacctgcccctcgggcgaataaaataaa 824
 Db 717 CCGTGAGGTTGTCACTAACTGCTCGGTTCGCCCAATAAATAAA 763

RESULT 13

BE733056 711 bp mRNA EST 15-SEP-2000
 LOCUS 601569557F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844119 5',
 DEFINITION mRNA sequence.

ACCESSION BE733056
 VERSION BE733056.1 GI:10147048
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: ATCC

REFERENCE

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLC540 row: a column: 16
 High quality sequence stop: 711.

FEATURES

source

1. 711
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3844119"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 116 a 246 c 227 g 122 t

ORIGIN

Query Match 78.7%; Score 648.2; DB 10; Length 711;
 Best Local Similarity 97.6%; Pred. NO. 1.5e-115;
 Matches 690; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 101 gctgcgcggttcccgataccttgggtctgctccagcgtggggcccggtgtgcaaggcac 160
 Db 7 GGCTGCGGCGCTCCCGATCCCTTGTGCTCGCTCCAGCGTGGCCGCGCTGTCAGGCAC 66
 QY 161 gagggttccatcagaagcgtggccaccgcatgcccgaagagacaccagcctgcccctgcaaa 220
 Db 67 GAGGTGTCCATCAGACGCTGGCCACCGATGGCCCAAGCAGCACCAGCTGCCTTGCCTTGC 126
 QY 221 agccagagagcgtgggtcccaaaaccagcagcgggggagtatgtgtggccaaagctgg 280
 Db 127 AGCCAGAGCCGTGGCTCCCAAAACCAGCAGCGGGGCGAGTATGTGTGSCCAGCTGG 186
 QY 281 atgaactctgtaactgggcccgcggaggttctctgtggcccatgaccttcgacctggcct 340
 Db 187 ATGACCTCTCAACTGGCGCCGCGGAGTCTCTGTGGCCCATGACCTTCGCGCTGGCCT 246
 QY 341 gctgcgccttgagatgatgcacatggagcaccgccctagacatgacacgctttggcg 400
 Db 247 GCTGCCCGTGGAGATGATGACATGGCAGCACCCTGCTACGACATGACCGCTTTGGCG 306
 QY 401 tggcttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 460
 Db 307 TGCTCTCCGCGCCAGCGCCGCGCAGTCCGACGTCATGATCGTGGCGCGCACACTACCA 366
 QY 461 acaagatggccccagcgtcttcgaaggtctacagacagatgccggagcggcgctactctgg 520
 Db 367 ACAAGATGGCCCCAGCGCTTCGCAAGGTCTACGACAGATGCCGAGCGCGCTACGTGG 426
 QY 521 tctccatggggagctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 580
 Db 427 TCTCCATGGGAGCTGCGCCCAACGAGGAGGCTACTACCACTATTCTCTACTCGGTGTGA 486
 QY 581 ggggtgcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 640
 Db 487 GGGGTGCGACCGCATCTGTCGCCGTGGACATCTAGATCCAGCTGCCACCTACAGCGCG 546
 QY 641 agccctctctacgcacatcccgagctgcagctgcagagaagatcaagcggagcggagctgc 700
 Db 547 AGCCCTCTCTACGCGCATCTGTCAGCTGCAGAGGAAGATCAAGCGGAGCGGAGCGCTTC 606
 QY 701 agatctgttacccgaggtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 760
 Db 607 AGATCTGTGACCGCAGGTAGCGCGCCGCCGCCGCCGCG-CGGAGCCTGTGCGCGTC--TGT 663
 QY 761 ccagcctctgtgtccctcagattgtcaataaacctgcctcgg 807
 Db 664 CCCAGCTGCTTGTGT-CCGTGAGGTTGTCAATAAACTGCCCTCGGG 709

RESULT 14

BI195871 788 bp mRNA EST 10-JUL-2001
 LOCUS 602756115F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4891946 5',
 DEFINITION mRNA sequence.
 ACCESSION BI195871
 VERSION BI195871.1 GI:14650891
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: ATCC

REFERENCE

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be


```
QY 312 tctgtgcccacatgaaccttcggcctggtgctgctgagccgtgagatgatgcacatggcagc 371
Db 241 TCCTGTGGCCCATGACCTTCGGCTGGCTGCTGCGCTGGAGATGATGCACATGGCAGC 300
QY 372 accccgctacgacatggaccgtttgggtgtgttccggccagcccgccagtcgca 431
Db 301 ACCCCGCTACGACATGGACCGTTTGGGTGTGTCTCCGCGCCAGCCGCGCCAGTCCGA 360
QY 432 cgtcatgatcgttggcgggcacactcaccaacaagaatggccccagcgcttcgcaaggtcta 491
Db 361 CGTCAATGATCGTGGCGGCACACTCACCAACAAGATGCCCCAGCGCTTCGCAAGGTCTA 420
QY 492 cgaccagatgcggagcgccgctacgtgtgtctccatggggagcctgcgccaaacgagagg 551
Db 421 CGACCAGATGCCGGAGCCGCGTACGTGTCTCCATGGGGAGCTGCCGCCAACGGAGGAGG 480
QY 552 ctactaccactattcctactcgttgggtggtgaggggtgcgacccgcacgtgcccgtggacat 611
Db 481 CTACTACCACATATTCCTACTCGTGGTGGAGGGCTGGCCAGCGCATCGTGCCTGGACAT 540
QY 612 ctacatcccaggctgcccacctacgcccagggccctgtctctacggcatcctgcagctgca 671
Db 541 CTACATCCCAGGCTG--CCACCTACGGC--GAGGCGTGTCTACGGCATCCTGSCAGCTGCA 597
QY 672 gaggaagatcaagcggagcgagcgtgcagatctggtaccgcaggtagcgccgcgccg 731
Db 598 GAGGAAGATCAAGCGGGAGCGGAGGCTGCAGATCTGGTA--CGCAGGTAG-----CGCGC 650
QY 732 cggcccgcgagcctgtgcgcctcctgtcctccagcctgtgttcccgtagagttgtca 791
Db 651 GCGCGCGGGAGCCTGTCTGCCCTCCTGTCCNCCAGCTGCTGTGTTCGCCGTGAGGTGTC 710
QY 792 ataaacctgccctcgggcaaaaaaaaaaaaaa 824
Db 711 ATRACCTGGCTCGCGGAAAAAATAAAAAA 743
```

Search completed: February 12, 2002, 17:41:38
Job time: 2873 sec

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1118	100.0	213	4	Q9BV17	Q9bv17 homo sapien
2	936.5	83.8	224	11	Q9DC70	Q9dc70 mus musculu
3	815.5	72.9	214	4	Q9HK5	Q9hk5 homo sapien
4	753.5	67.4	221	5	Q9VXK7	Q9vxk7 drosophila
5	743	66.5	212	5	Q9VAK5	Q9vak5 drosophila
6	731	65.4	213	10	Q9SP38	Q9sp38 lupinus lut
7	721.5	64.5	192	10	Q9LK98	Q9lk98 lupinus lut
8	716	64.0	158	8	Q9GRU4	Q9gru4 rhodomonas
9	712.5	63.7	210	10	Q9LKH4	Q9lkh4 lupinus lut
10	702	62.8	210	3	Q9UUT7	Q9uut7 yarrowia li
11	699	62.5	158	8	Q9TCA4	Q9tca4 nephroselmi
12	654	58.5	191	2	Q9PGJ4	Q9pgj4 xyella fas
13	642.5	57.5	193	2	Q9A6X1	Q9a6x1 caulobacter
14	613	54.8	160	2	Q9K1C2	Q9k1c2 neisseria m
15	605	54.1	160	2	Q9JX79	Q9jx79 neisseria m
16	598	53.5	162	8	Q9XMU2	Q9xmu2 tetrahymena
17	497.5	44.5	159	2	Q25851	Q25851 helicobacte
18	496.5	44.4	159	2	Q9ZJW6	Q9zjw6 helicobacte
19	474	42.4	181	2	Q9RU87	Q9ru87 deinococcus

DE	MV017 PROTEIN.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RI	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=BRAIN;
RC	
RA	Mao Y.M., Xie Y., Huang X.Y., Ying K., Dai J.L.;
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF060512; AAC43130.1;
DR	InterPro: IPR002096; Complex1_20kD.
DR	Pfam: PF01058; oxidored_g6_1_
SQ	SEQUENCE 214 AA; 23622 MW; 8DC57E13F2FC07E3 CRC64;
Query Match 72.9%; Score 815.5; DB 4; Length 214;	
Best Local Similarity 75.3%; Pred. No. 3.7e-68;	
Matches 171; Conservative 7; Mismatches 22; Indels 27; Gaps	
QY	1 MAVLSAPGLRGFRILGRSSVGPAVGARGVHQSVATDGGSPSTQPALPKARAVAPKPSRG 60
Db	
Db	1 MAVLSAPGLRGFRILGRSSVGLAVGARGVHQSVATDGGSPSTQPALPKARAVAPKPSRG 60
QY	61 EYVVAKLDDLNNARRSSLWPMTFGACCACAVEMHMAAPRYMDRGVVFRASPROSDVM 120
Db	
Db	61 EYVVAKLDDLNNARRSSLWPMTFGACCACAVEMHMAAPRYMDRGVVFRASPROSDVM 120
QY	121 IVAGTLTNKMAPALKRVYDOMPEPVVYVNGSANGGGYYHYHSYSVVRGCDRIVPVDI-- 178
Db	
Db	121 IVAGTLTNKMAPALKR---SPRRP-----DAGAALRGHLGRQLRGRLLPLFLIG 167
QY	179 -----YIPCPPTAEALLYGILOQRKKRRERLQIWYRR 213
Db	
Db	168 GEGLRPHRHGHLPRLPTAEALLYGILOQRKKRRERLQIWYRR 214
RESULT 4	
Q9VAXK7	PRELIMINARY; PRT; 221 AA.
ID Q9VXK7	
AC Q9VXK7	
DT 01-MAY-2000 (TREMBLrel. 13, Created)	
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE CG1172 PROTEIN.	
GN C9N172.	
OS Drosophila melanogaster (Fruit fly).	
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC Prteryozoa; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC Ephydroidea; trosophilidae; Drosophilla.	
ON NCBI_Taxid=7227;	
OX [1]	
RC SEQUENCE FROM N.A.	
RP STRAIN=BERKELEY;	
RC MEDLINE=20196006; PubMed=10731132;	
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,	
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,	
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,	
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA Abril J.F., Acbayani A., An H.-J., Andrews-pfankoch C., Baldwin D.,	
RA Ballaw R.M., Basu A., Bavendale J., Bayraktaroglu L., Beasley E.M.,	
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandardi D., Bolshakov S.,	
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,	
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,	
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,	


```

PC TISSUE-ROOT;
RA Piekna D., Sikorski M., Augustyniak H.;
RT "L. luteus mRNA for PSST subunit of the NADH:ubiquinone oxidoreductase
RL complex.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF185271; AAF01037.1; -.
DR InterPro: IPR002096; Complex1_20KD.
DR Pfam: PF01058; Oxidored_g6; 1.
DR PROSITE: PS01150; COMPLEX1_20K; 1.
KW Ubiquinone.
SQ SEQUENCE 213 AA; 23608 MW; CB39C11B002215E8 CRC64;

Query Match 65.4%; Score 731; DB 10; Length 213;
Best Local Similarity 63.2%; Pred. No. 2.7e-60;
Matches 139; Conservative 28; Mismatches 39; Indels 14; Gaps 4;

QY 1 MAYLSAPGLR-----GFRILGRSSGPAQARGVHQSVATDGPSSQTQPALEKARAVA 53
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 1 MALLTRATRLPOLAPSSHRILSLHTL-PSLSP----DTAATPAPYSRPP--PPSTSSST 53
QY 54 KPSSRGVYVAKLDDLNVNARRSSLPMTFGLACCAVEMHMAAPRYDMDRFGVVFRA 113
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 54 AGISKAETAEFVSKVDLNMWARRGSIWMTFGLACCAVEMHMTGAARYDLDLRFGIIFPS 113
QY 114 PROSDYIVAGTLTNKMAPALKRVYDQPEPRYVYVSMGSCANGGGYHYHSYVVRGCDRI 173
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 114 PROSDCMIVAGTLTNKMAPALKRVYDQPEPRYVYVSMGSCANGGGYHYHSYVVRGCDRI 173
QY 174 VPVDIIPGCPPTAEALLYILOLQKIKRRRLQIWR 213
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 174 VPVDIYVPGCPPTAEALLYILOLQKIKRRNRKDFLLMWT 213

RESULT 7
Q9LKG9 PRELIMINARY; PRT; 192 AA.
AC Q9LKG9;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT PSST (FRAGMENT).
OS Lupinus luteus (yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RA Piekna D., Sikorski M., Augustyniak H.;
RT "L. luteus mRNA for PSST subunit of the NADH:ubiquinone
RL oxidoreductase.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF281035; AAF91434.1; -.
DR InterPro: IPR002096; Complex1_20KD.
DR Pfam: PF01058; Oxidored_g6; 1.
DR PROSITE: PS01150; COMPLEX1_20K; 1.
KW Ubiquinone.
FT NON-TER
SQ SEQUENCE 192 AA; 21418 MW; 02B1B848B27104C5 CRC64;

Query Match 64.5%; Score 721.5; DB 10; Length 192;
Best Local Similarity 71.0%; Pred. No. 1.8e-59;
Matches 132; Conservative 20; Mismatches 21; Indels 13; Gaps 2;

QY 41 STQPAL-----PKARAVAPKPS-----RGEYVYVAKLDDLNVNARRSSLPMTFGLA 87
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 7 TTLPSLSPETATTPARPPPTSSPAGISKAETAEVSKVDLNMWARRGSIWMTFGLA 66
QY 88 CCAVEMHMAAPRYDMDRFGVVFRAASPRQSDYVMTVAGTLTNKMAPALKRVYDQPEPRYV 147
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 67 CCAVEMHMTGAARYDLDLRFGIIFRPSRQSDCMIVAGTLTNKMAPALKRVYDQPEPRYV 126

```

```

QY 148 VSMGSCANGGGYHYHSYVVRGCDRIYVVDIYVPGCPPTAEALLYILOLQKIKRRRL 207
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 127 ISMGSCANGGGYHYHSYVVRGCDRIYVVDIYVPGCPPTAEALLYILOLQKIKRRKDF 186
QY 208 QTWYRR 213
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 187 LLMWTK 192

RESULT 8
Q9G8U4 PRELIMINARY; PRT; 158 AA.
ID Q9G8U4;
AC Q9G8U4;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 10 (EC 1.6.5.3).
GN NAD10.
OS Rhodomonas salina.
OC Rhodomonadion.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Rhodomonas.
OX NCBI_TaxID=52370;
RN [1]
RP SEQUENCE FROM N.A.
RA Burger G., Lang B.F., Maier U.G., McFadden G.I., Gray M.W.;
RT "Algae with secondary chloroplasts have mitochondria that originate
RT from the host.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF288090; AAG17753.1; -.
DR InterPro: IPR002096; Complex1_20KD.
DR Pfam: PF01058; Oxidored_g6; 1.
DR PROSITE: PS01150; COMPLEX1_20K; 1.
KW Oxidoreductase; Mitochondrion.
SQ SEQUENCE 158 AA; 17819 MW; 0798BB29D71611C4 CRC64;

Query Match 64.0%; Score 716; DB 8; Length 158;
Best Local Similarity 79.9%; Pred. No. 4.6e-59;
Matches 123; Conservative 22; Mismatches 9; Indels 0; Gaps 0;

QY 58 SRGEYVYVAKLDDLNVNARRSSLPMTFGLACCAVEMHMAAPRYDMDRFGVVFRA 117
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 3 NNTKVFISKIDLLNWARSSLPMTFGLACCAVEMHMAAGASRYDFDFRFGIIFRA 62
QY 118 DVMIVAGTLTNKMAPALKRVYDQPEPRYVYVSMGSCANGGGYHYHSYVVRGCDRIYVVD 177
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 63 DIMIVAGTLTNKMAPALKRVYDQPEPRYVYVSMGSCANGGGYHYHSYVVRGCDRIYVVD 122
QY 178 IYIPGCPPTAEALLYILOLQKIKRRRLQIY 211
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 123 VFVPGCPPTAEALLYILOLQKIKRRTKVTQTF 156

RESULT 9
Q9LKH4 PRELIMINARY; PRT; 210 AA.
ID Q9LKH4;
AC Q9LKH4;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT PSST.
OS Lupinus luteus (yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RA Piekna D., Sikorski M., Augustyniak H.;
RT "Lupinus luteus PSST gene.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF279902; AAF91429.1; -.

```


OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=9A5C;
 RC MEDLINE=20365717; PubMed=10910347;
 RX SIMPSON A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Canaro A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohenisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.P., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silveira M.L.E., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003884; AAF83117.1;
 DR InterPro; IPR002096; Complexl_20KD.
 DR Pfam; PF01058; oxidored_q6; 1.
 DR PROSITE; PS01150; COMPLEXL_20K; 1.
 KW Complete proteome.
 SQ SEQUENCE 191 AA; 20858 MW; 7A14CC80A8857CF5 CRC64;

Query Match 58.5%; Score 654; DB 2; Length 191;
 Best Local Similarity 72.1%; Pred. No. 3.4e-53;
 Matches 119; Conservative 14; Mismatches 20; Indels 12; Gaps 1;
 QY 39 PSSTOPALPKARAVAPKPSRGEYVYVAKLDLVNKAARSLWPMTFLGACCAVEMHMAA 98
 Db 33 POGESPLQKG-----YVTSVDALLNWTGSMWPTFLGACCAVEMHMAA 80
 QY 99 PRYDMDFGVWFRASPRQSDVMIVAGTLTNKMAPALRKVDQMPPEPRYVSMGSCANGGG 158
 Db 81 ARILDRLGVFRPSRQSDVMIVAGTLVKNMAPALRKVDQMPDPKRWISMGSCANGGG 140
 QY 159 YVHSYVVRGCDRVPVDIYVPGCPPTAEALLYGILQLOKKIR 203
 Db 141 YVHSYVVRGCDRVPVDIYVPGCPPTAEALYVGILOKKIWR 185

RESULT 13
 Q9A6X1
 ID Q9A6X1 PRELIMINARY; PRT; 193 AA.
 AC Q9A6X1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NADH DEHYDROGENASE I, B SUBUNIT.
 GN CC1955.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=2371;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=9A5C;
 RC MEDLINE=20365717; PubMed=10910347;
 RX SIMPSON A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Canaro A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohenisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.P., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silveira M.L.E., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003884; AAF83117.1;
 DR InterPro; IPR002096; Complexl_20KD.
 DR Pfam; PF01058; oxidored_q6; 1.
 DR PROSITE; PS01150; COMPLEXL_20K; 1.
 KW Complete proteome.
 SQ SEQUENCE 191 AA; 20858 MW; 7A14CC80A8857CF5 CRC64;

OX NCBI_TaxID=69394;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21173698; PubMed=11259647;
 RX Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005868; AAK23930.1;
 DR TIGR; CC1955;
 KW Complete proteome.
 SQ SEQUENCE 193 AA; 21082 MW; C5F57AE6B341862 CRC64;
 Query Match 57.5%; Score 642.5; DB 2; Length 193;
 Best Local Similarity 66.1%; Pred. No. 4e-52;
 Matches 123; Conservative 20; Mismatches 26; Indels 17; Gaps 2;
 QY 34 VAIDGPSSTQPALPKARA-----VAPKPSRGEYVYVAKLDLVNWAARS 77
 Db 3 VIVGNSSPVPAISAGRTVEGYDPKLDHPDFDGVSOQLADKG-FITAAADDLITWARTG 61
 QY 78 SLWPTFGLACCAVEMHMAAPRYDMDFGVWFRASPRQSDVMIVAGTLTNKMAPALRKV 137
 Db 62 SLWPTFGLACCAVEMHMAAPRYDMDFGVWFRASPRQSDVMIVAGTLTNKMAPALRKV 121
 QY 138 YQDMPPEPRYVSMGSCANGGGYVHSYVVRGCDRVPVDIYVPGCPPTAEALLYGILQL 197
 Db 122 YQDMPPEPRYVSMGSCANGGGYVHSYVVRGCDRVPVDIYVPGCPPTAEALYVGIQL 181
 QY 198 QKKIR 203
 Db 182 QKKIR 187
 RESULT 14
 Q9K1C2
 ID Q9K1C2 PRELIMINARY; PRT; 160 AA.
 AC Q9K1C2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE NADH DEHYDROGENASE I, B SUBUNIT.
 GN NMG0242.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002381; AAF40696.1;
 DR TIGR; NMG0242;
 DR InterPro; IPR002096; Complexl_20kD.
 DR Pfam; PF01058; oxidored_q6; 1.
 DR PROSITE; PS01150; COMPLEXL_20K; 1.
 KW Complete proteome.


```
SQ SEQUENCE 160 AA; 17629 MW; F1616423A8D799AF CRC64;

Query Match      54.1%; Score 613; DB 2; Length 160;
Best Local Similarity 73.9%; Pred. No. 1.8e-49;
Matches 105; Conservative 22; Mismatches 15; Indels 0; Gaps 0;

QY 62 YVAKLDLVNWARSSLPWTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQSDVMI 121
Db 11 FITTSADTVLNMRTGSLWPTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQADLMI 70
QY 122 VAGTLTNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGCDRIYVPDIYIP 181
Db 71 VAGTLTNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGADRVVPDVYVP 130
QY 182 GCPPTAEALYGLIQLQKIRK 203
Db 131 GCPPTAEALYGLIQLQKIRK 152

RESULT 15
Q9JX79
ID Q9JX79 PRELIMINARY; PRT; 160 AA.
AC Q9JX79;
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)
DE NADH DEHYDROGENASE I CHAIN B (EC 1.6.5.3).
GN NUOB OR NMA0018.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65569;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162752; CAB83338.1;
DR InterPro; IPR002096; Complex1_20KD.
DR Pfam; PF01058; oxidored_g6; 1.
DR PROSITE; PS01150; COMPLEX1_20K; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 160 AA; 17619 MW; 4E9CF64C554799AC CRC64;

Query Match      54.1%; Score 605; DB 2; Length 160;
Best Local Similarity 73.2%; Pred. No. 9.9e-49;
Matches 104; Conservative 22; Mismatches 16; Indels 0; Gaps 0;

QY 62 YVAKLDLVNWARSSLPWTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQSDVMI 121
Db 11 FITTSADTVLNMRTGSLWPTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQADLMI 70
QY 122 VAGTLTNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGCDRIYVPDIYIP 181
Db 71 VAGTLTNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGADRVVPDVYVP 130
QY 182 GCPPTAEALYGLIQLQKIRK 203
Db 131 GCPPTAEALYGLIQLQKIRK 152
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 19:56:10 ; Search time 37.72 Seconds
(without alignments)
418.282 Million cell updates/sec

Title: US-09-525-867-1

Perfect score: 1118

Sequence: 1 MAVLSAPGLRGFRILGLRSS.....ILQLQKIKRRRLQIWR 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1118	100.0	213	AA39624	Human polypeptide
2	1118	100.0	214	AA39624	Human polypeptide
3	826	73.9	175	AA34291	Human secreted pro
4	813	72.7	171	AA34292	Human secreted pro
5	731.5	65.4	200	AA33025	Zea mays protein f
6	725	64.8	216	AA34738	Arabidopsis thalia
7	725	64.8	216	AA34738	Arabidopsis thalia
8	725	64.8	218	AA34737	Arabidopsis thalia
9	725	64.8	218	AA34736	Arabidopsis thalia
10	705	63.1	158	AA33026	Zea mays protein f
11	663	59.3	143	AA33027	Zea mays protein f

12	661	59.1	143	21	AA34739	Arabidopsis thalia
13	661	59.1	143	21	AA34739	Arabidopsis thalia
14	652	58.3	143	21	AA34739	Arabidopsis thalia
15	602	53.8	132	21	AA34739	Arabidopsis thalia
16	338.5	30.3	132	21	AA34739	Arabidopsis thalia
17	88.5	7.9	510	19	AA34739	Arabidopsis thalia
18	82	7.3	729	21	AA34739	Arabidopsis thalia
19	79.5	7.1	319	15	AA34739	Arabidopsis thalia
20	79.5	7.1	319	15	AA34739	Arabidopsis thalia
21	78	7.0	892	22	AA34739	Arabidopsis thalia
22	77.5	6.9	361	22	AA34739	Arabidopsis thalia
23	77.5	6.9	361	22	AA34739	Arabidopsis thalia
24	77.5	6.9	361	22	AA34739	Arabidopsis thalia
25	77.5	6.9	361	22	AA34739	Arabidopsis thalia
26	77.5	6.9	361	22	AA34739	Arabidopsis thalia
27	77.5	6.9	361	22	AA34739	Arabidopsis thalia
28	77.5	6.9	361	22	AA34739	Arabidopsis thalia
29	77	6.9	257	22	AA34739	Arabidopsis thalia
30	76.5	6.8	753	22	AA34739	Arabidopsis thalia
31	76	6.8	369	21	AA34739	Arabidopsis thalia
32	76	6.8	374	21	AA34739	Arabidopsis thalia
33	76	6.8	412	21	AA34739	Arabidopsis thalia
34	76	6.8	566	22	AA34739	Arabidopsis thalia
35	76	6.8	635	22	AA34739	Arabidopsis thalia
36	76	6.8	660	22	AA34739	Arabidopsis thalia
37	75.5	6.8	347	19	AA34739	Arabidopsis thalia
38	75.5	6.8	548	19	AA34739	Arabidopsis thalia
39	75.5	6.8	999	21	AA34739	Arabidopsis thalia
40	75.5	6.8	1008	21	AA34739	Arabidopsis thalia
41	75.5	6.8	1065	21	AA34739	Arabidopsis thalia
42	75.5	6.8	3011	16	AA34739	Arabidopsis thalia
43	74.5	6.7	164	22	AA34739	Arabidopsis thalia
44	74.5	6.7	500	21	AA34739	Arabidopsis thalia
45	74.5	6.7	500	21	AA34739	Arabidopsis thalia

ALIGNMENTS

RESULT	1
AA39624	AA39624 standard; Protein: 213 AA.
ID	AA39624 standard; Protein: 213 AA.
XX	XX
AC	AA39624;
XX	XX
DT	22-OCT-2001 (first entry)
XX	XX
DE	Human polypeptide SEQ ID NO 2769.
XX	XX
KW	Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
OS	Homo sapiens.
XX	XX
PN	WO200153312-A1.
PD	26-JUL-2001.
XX	XX
PF	26-DEC-2000; 2000WO-US34263.
XX	XX
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	XX

(HYSE-) HYSEQ INC.

PA XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA158780.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 4; SEQ ID NO 2769; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC assays for receptor activity, cancer diagnosis and therapy, drug screening,
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 213 AA;

Query Match 100.0%; Score 1118; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 5.1e-117;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLSAPGLRGFRILGRSSVGPVQARGVHOSVATDGPSSQTQALPKARAVAPKPSRG 60
 Db 1 mavlsapglrgfrilgrssvqavqargvhqsvatdgpssqtqalpkaravapkpssrg 60

QY 61 EYVAKLDDLNVNARRSSLWPMTFGLACCAVEMHMAAPRYDMDRFGVVFRRASPRQSDVM 120
 Db 61 eyvvaklddlvnvarrsslwpmftfglaccavemhmaprydmdrfgvvfrasprqsdvm 120

QY 121 IVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIYPVDIYI 180
 Db 121 ivagtltnkmapalrkvydqmpeprvyvsmgscanggggyhyhsyvvrgcdriypvdiyi 180

QY 181 PGCPTAEALLYGILQKIKRRRLQIYYRR 213
 Db 181 pgcptaeallygilqlqkikrrrlqiwyrr 213

RESULT 2
 AAM41410
 ID AAM41410 standard; Protein; 214 AA.

XX AC AAM41410;
 XX DT 22-OCT-2001 (first entry)
 XX Human polypeptide SEQ ID NO 6341.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.

XX WO200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.
 XX 19-OCT-2000; 2000US-0693036.
 XX 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA160566.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 6341; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC assays for receptor activity, cancer diagnosis and therapy, drug screening,
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 214 AA;

Query Match 100.0%; Score 1118; DB 22; Length 214;
 Best Local Similarity 100.0%; Pred. No. 5.2e-117;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLSAPGLRGFRILGRSSVGPVQARGVHOSVATDGPSSQTQALPKARAVAPKPSRG 60
 Db 2 mavlsapglrgfrilgrssvqavqargvhqsvatdgpssqtqalpkaravapkpssrg 61

QY 61 EYVAKLDDLNVNARRSSLWPMTFGLACCAVEMHMAAPRYDMDRFGVVFRRASPRQSDVM 120
 Db 62 eyvvaklddlvnvarrsslwpmftfglaccavemhmaprydmdrfgvvfrasprqsdvm 121

QY 121 IVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIYPVDIYI 180
 Db 122 ivagtltnkmapalrkvydqmpeprvyvsmgscanggggyhyhsyvvrgcdriypvdiyi 181

QY 181 PGCPTAEALLYGILQKIKRRRLQIYYRR 213
 Db 182 pgcptaeallygilqlqkikrrrlqiwyrr 214

RESULT 3
 AAB34291

ID AAB34291 standard; Protein; 175 AA.
 AC AAB34291;
 XX
 XX
 DT 02-FEB-2001 (first entry)
 DE Human secreted protein BLAST search protein SEQ ID NO: 137.
 XX
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200005352-A2.
 PN
 XX
 XX 21-SEP-2000.
 PD
 XX
 XX 09-MAR-2000; 2000WO-US06044.
 PF
 XX
 XX 12-MAR-1999; 99US-0124099.
 PR
 XX 03-DEC-1999; 99US-0168661.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI
 XX
 XX WPI; 2000-602124/57.
 DR
 XX
 XX Novel human secreted proteins useful for diagnosis, prevention and
 PT treatment of disorders including neurological, cell proliferative,
 PT cardiovascular, autoimmune and inflammatory disorders and microbial
 PT infections
 XX
 PS Disclosure; Pages 379-380; 383pp; English.
 XX
 CC The invention relates to the isolation of genes AAC59507-C59556 encoding
 CC the human secreted proteins AAB34218-B34264. This sequence represents a
 CC peptide fragment homologous to the protein encoded by the gene isolated
 CC in the present invention. The sequence is a search result from a BLASTX
 CC homology search. The genes and proteins are useful for preventing,
 CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
 CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases
 CC as e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
 CC as viral, bacterial, fungal and parasitic infections.
 XX
 XX Sequence 175 AA;
 SQ
 Query Match 73.9%; Score 826; DB 21; Length 175;
 Best Local Similarity 88.6%; Pred. No. 2e-84; Indels 4; Gaps 1;
 Matches 155; Conservative 7; Mismatches 9;
 QY 14 IGLRSSVGPVQARGVHQSVATDGPSTQPALPKARAVAPK----PSSRGYYVAKLDD 69
 DB 1 ILAVRSGVGAALQVRGSHSMAADSPSTQPAVSGATAVVPKPAALPSSRGYYVAKLDD 60
 QY 70 LYNWARRSLWPTGLACCAVMMHMAAPRYDMDRGVYFRASPRSDVMIVAGILTNN 129
 DB 61 ILNARRSLWPTGLACCAVMMHMAAPRYDMDRGVYFRASPRSDVMIVAGILTNN 120
 QY 130 MAPALRKVDQMPPEPRYYVSMGSCANGGGYYHSYSVVRGCDRIVPVDIYIPGCP 184

DB 121 MAPALRKVDQMPPEPRYYVSMGSCANGGGYYHSYSVVRGCDRIVPVDIYIPGCP 175
 RESULT 4
 AAB34292
 ID AAB34292 standard; Protein; 171 AA.
 XX
 AC AAB34292;
 XX
 DT 02-FEB-2001 (first entry)
 DE Human secreted protein BLAST search protein SEQ ID NO: 138.
 XX
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200005352-A2.
 PN
 XX
 XX 21-SEP-2000.
 PD
 XX
 XX 09-MAR-2000; 2000WO-US06044.
 PF
 XX
 XX 12-MAR-1999; 99US-0124099.
 PR
 XX 03-DEC-1999; 99US-0168661.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI
 XX
 XX WPI; 2000-602124/57.
 DR
 XX
 XX Novel human secreted proteins useful for diagnosis, prevention and
 PT treatment of disorders including neurological, cell proliferative,
 PT cardiovascular, autoimmune and inflammatory disorders and microbial
 PT infections
 XX
 PS Disclosure; Page 382; 383pp; English.
 XX
 CC The invention relates to the isolation of genes AAC59507-C59556 encoding
 CC the human secreted proteins AAB34218-B34264. This sequence represents a
 CC fragment of the protein encoded by the gene isolated in the present
 CC invention. The sequence is used as a query sequence for doing BLASTX
 CC searches to determine homologous sequence to the protein. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and
 CC ulcerative colitis; (c) cardiovascular disorders such as myocardial
 CC ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral
 CC anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections.
 XX
 XX Sequence 171 AA;
 SQ
 Query Match 72.7%; Score 813; DB 21; Length 171;
 Best Local Similarity 91.8%; Pred. No. 5.5e-83;
 Matches 157; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 14 IGLRSSVGPVQARGVHQSVATDGPSTQPALPKARAVAPKPSRGYYVAKLDDLVNN 73
 DB 1 ILAVRSGVGAALQVRGSHSMAADSPSTQPAVSGATAVVPKPAALPSSRGYYVAKLDD 60

```
Db 1 ilgirssgvpaqargvhqsvatdgsstqpalpkaravapkpssrggyvvakldlvnw 60
QY 74 ARRSLLPMTFGLACCAVENMMAAPRYDMDFGVVFRASPRQSDVMIVAGTLTKMAPA 133
    |||||||
Db 61 arrrsllpmtfglaccavemhnaaprydmndrfgvvfraspqsdvmivagtltnkmapa 120
    |||||||
QY 134 LRKYVDQMPPEPRYVWSMGSCANGGYYHYHSYSVVRCGDRIVPVDIYIPGCP 184
    |||||||
Db 121 lrkyvdqmppepryvvsmgscanxxxxxxxccdrivpvdlyipgcp 171

RESULT 5
AAG33025
ID AAG33025 standard; Protein; 200 AA.
XX AC AAG33025;
XX XX
DT 17-OCT-2000 (first entry)
XX XX
DE Zea mays protein fragment SEQ ID NO: 39954.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX XX
OS Zea mays subsp. mays.
XX XX
FN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 05-MAR-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 23-MAR-1999; 99US-0123548.
PR 25-MAR-1999; 99US-0123788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 30-APR-1999; 99US-0130891.
PR 04-MAY-1999; 99US-0131449.
PR 05-MAY-1999; 99US-0132048.
PR 06-MAY-1999; 99US-0132407.
PR 07-MAY-1999; 99US-0132485.
PR 11-MAY-1999; 99US-0132486.
PR 14-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 09-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```

```
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160788.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 65.4%; Score 731.5; DB 21; Length 200;
Best Local Similarity 73.2%; Pred No. 9.1e-74;
Matches 134; Conservative 17; Mismatches 31; Indels 1; Gaps 1;

QY 31 HQSVATDGPSTQPALPKARAVAPKPSRGEYVYVAKLDLNVNARRSSLWPMTFGLACCA 90

Db 19 ysaaaaasptspa-pyggapppamstaeafvyskvddlmwarksgiswpmntfglacc 77
QY 91 VEMHMAAPRYDMDRFVGVFRASPRQSDVMIVAGTLTNKMAPALRKVYDQMPEPRVYSM 150
Db 78 vemhbagasrydfdrfgvlfprsprqscdmivaglttnkmapalrkvydqmpewrvism 137
QY 151 GSCANGGGYHYHSYSVVRGCDRIYVVDIYIPGCPPTAEALLYGILQORKIKRERLQIW 210
Db 138 gscangggynhsyvsrvrgcdriyvpdiyvpogpptaallvgvlqkknrrkdfhw 197
QY 211 YRR 213
Db 198 WTK 200

RESULT 6
AAG14738
ID AAG14738 standard; Protein; 216 AA.
XX
AC AAG14738;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14714.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
```

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

[illegible]

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

4

```
PR 22-JUL-1999; 99US-01451192.
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-01452118.
PR 23-JUL-1999; 99US-01452118.
PR 26-JUL-1999; 99US-01452224.
PR 26-JUL-1999; 99US-01452276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147132.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149388.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159234.
PR 13-OCT-1999; 99US-0159235.
PR 14-OCT-1999; 99US-0159323.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.8%; Score 725; DB 21; Length 218;
Best Local Similarity 68.2%; Pred. No. 5.5e-73;
Matches 135; Conservative 20; Mismatches 35; Indels 8; Gaps 2;

QY 24 AVQAGV---HQSVDGPPS-----TOPALPKARAVAPKPSRGGEYVAKLDOLYNWAR 75
DB 21 avaaasvhlhtslpalspsstsytrpgpptspppglskaaeviskvdldnnwar 80
QY 76 RSSLWPMTEGLACCAVEMHMAAPRYDMDFGVVFRASPRQSDVMIVAGLTNKNMAPALR 135
DB 81 tgsiwpmtcflaccavemhmtgaarydlrfgiifrsprqscdmivagltitnkmapai 140
QY 136 KYVDQMPERYVVMGSCANGGYHYHSYVVRGCDRIVPDYIYPCGPTAEALLYGIL 195
DB 141 kvydgmpeprwvismgscanggyhyhsyvsrgcdripvdiyvpcgptaeallygl 200
QY 196 QLQRKIKRERRLQIYWR 213
DB 201 qlqkkinrrkflhwnk 218

RESULT 10
AAG33026
ID AAG33026 standard; Protein: 158 AA.
XX
AC AAG33026;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 39955.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
```

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132486.
PR 14-MAY-1999; 99US-0132418.
PR 14-MAY-1999; 99US-0132419.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

[illegible]

```
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159639.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 59.3%; Score 663; DB 21; Length 143;
Best Local Similarity 81.8%; Pred. NO. 2.7e-66;
Matches 117; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
QY 71 VNWARSSLPMTFTGLACCAVEMHMAAPRYDMDFGVYFRASPRQSDVMIVAGTLTNKM 130
Db 1 mnwarkgsiwpmtfglaccavemmhagasydfdfgvrpsrqscmivagtltnkm 60
QY 131 APALRKVDMPPEPRVYVSMGANGGGYHYHSYVVRGCDRIVPDYIPGCPPTAEAL 190
Db 61 apalrkvydmppeprvismgscanggggyhyhsyvvrgcdriypdviyvgpcptaeal 120
QY 191 LYGILOLQKIKRRRLQIWR 213
Db 121 lygvlqlqkikrrkdfhlhwtk 143
RESULT 12
AAG14739
ID AAG14739 standard; Protein; 143 AA.
XX
AC AAG14739;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14715.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
```


PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136031.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0130884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149388.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

```

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.1%; Score 661; DB 21; Length 143;
Best Local Similarity 81.1%; Pred. No. 4.5e-66;
Matches 116; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

Qy 71 VNWARRSLPMTFGLACCAVENMHMAAPRYDMDFGVFRASPRQSDVMIVAGTLTNKM 130
Db 1 mnwartsipmtfglaccavemmhgaaaryldirgfrpsrqscdmvagtctkm 60

Qy 131 APALRKVDMPPEPVYVMSGANGGYHYHSYSVVRGCDRVPVDIYIPGCPPTAEAL 190
Db 61 apalrkvydmppepvysngscanggyhyhsysvvrsgcdripvdiypgcpptaeal 120

Qy 191 LYGILQLOKIKRERLQIWR 213
Db 121 lygllqkkinrkdfhlwnk 143

RESULT 13
AAG18238
ID AAG18238 standard; Protein; 143 AA.
XX
XX AAG18238;
AC
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 19565.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.

```

```

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

```

PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 28-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.1%; Score 661; DB 21; Length 143;
Best Local Similarity 81.1%; Pred. No. 4.5e-66;
Matches 116; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 71 VNWARRSLWPTFGLACCAVENMHMAAPRYDMRQGVVFRASPRQSDVNIYAGLTNKM 130
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 mnwartgsiwpmtfglaccavemhmtgaarydlidrfgi:frpsprqsdcmivagltlnkm 60
QY 131 APALRKVYDQMPPEPRVYVSMGSCANGGYYHYSYVVRGCDRIVPDVYIPGCPPTAEAL 190
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 apalrkvydqmpeprvismgscangggyyhysvvrqcdriypdvdyvpgcpptaeal 120
QY 191 LYGILQLRKIKRRLQIWR 213
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 lygilqlqkkinrrkdfhlwnk 143

RESULT 14

AAG28057
ID AAG28057 standard; Protein; 143 AA.

XX AC AAG28057;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33132.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130591.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 08-MAY-1999; 99US-0132486.
PR 08-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144008.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154033.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.